

# Product variety and economic growth : the counteracting effects of scale and idiosyncrasy

Citation for published version (APA):

Straathof, S. M. (2005). *Product variety and economic growth : the counteracting effects of scale and idiosyncrasy*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20050429ss>

## Document status and date:

Published: 01/01/2005

## DOI:

[10.26481/dis.20050429ss](https://doi.org/10.26481/dis.20050429ss)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

## PRODUCT VARIETY AND ECONOMIC GROWTH

© 2005 S.M. Straathof  
All rights reserved.

Cover design by Marloes Boutkan  
Typeset in Computer Modern  
Prepared with  $\text{\LaTeX}$ , using the `memoir` class

Published by Universitaire Pers Maastricht  
ISBN 90 5278 457 4  
Printed in the Netherlands by Datawyse Maastricht

# Product variety and economic growth

The counteracting effects of scale and idiosyncrasy

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,  
op gezag van de Rector Magnificus, Prof. Mr. G.P.M.F. Mols,  
volgens het besluit van het College van Decanen,  
in het openbaar te verdedigen op  
vrijdag 29 april 2005 om 16.00 uur

door

Sebastiaan Michael Straathof







# Acknowledgements

This is the place where one ought to complain about the hardship endured while writing the thesis and where one ought to thank those people that supported you, as well as to apologize to all that suffered in the process. I will do all this shortly, but first I want to mention something else. Writing a thesis, and especially doing the research on which it is founded, essentially is a fun thing to do. Few people get (and take) the opportunity to spend four years on inquiry and discovery. I thank Adriaan van Zon and Bart Verspagen for giving me those years - and letting me exploit them thoroughly!

When I started there in October 2000, MERIT was a quiet place. I shared a room with Franco Bevilacqua, a supporter of Juve from Milan, who gave me insight into the peculiarities and logic of Italy and, indirectly, into those of The Netherlands. Later, Franco's place was briefly taken over by Martin Strobel, and after that by Sabine Fuss, who provided a steady supply of presents and coffee. Thank you all for your good company. At the moment, MERIT no longer is the quiet place that it was when I arrived. Thanks to Semih Akcomak, Abraham Garcia, Ekin Keskin, Tobias Kronenberg, Saeed Parto, and plenty of Italian and French visitors, this has changed substantially.

My PhD position has been part of the TEG-project and I benefitted from the critical and helpful comments that I got at several TEG-workshops. Thanks Bart van Ark, Jojo Jacob, Bart Los, Christoph Meister, Eddy Szirmai, and Marcel Timmer. I also have to thank all those at the University of Maastricht and elsewhere who generously shared some of their knowledge with me: Bert Balk, Lex Borghans, Robin Cowan, Erwin Diewert, Arnaud Dupuy, Michael Funke, Elad Harison, Knick Harley, Hugo Hollanders, Peter Howitt, Huub Meijers, Pierre Mohnen, Bulat Sanditov, Holger Strulik, Manuel Trajtenberg, Bas ter Weel (many times!), Klaus Wälde, Thomas Ziesemer, and all those I forget to mention here. The members of the reading committee gave me some useful comments that put the finishing touch to what is lying here before you. Wilma Coenegrachts, Silvana de Sanctis, and Corien Gijsbers were always ready to help me out with smaller and larger problems.

A PhD is rarely suffered alone and I am grateful to Marloes for her moral and practical support (ice lollies!) as well as her patience.

Bas Straathof  
Maastricht, March 2005



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Variety is the spice of life . . . . .	1
1.2	One assumption and two propositions . . . . .	2
1.2.1	New goods and the variety of consumer products . . . . .	2
1.2.2	Product variety and the size of the economy . . . . .	4
1.2.3	Product variety and idiosyncratic firm performance . . . . .	5
1.3	Research outline . . . . .	6
	Notes . . . . .	9
<b>2</b>	<b>Product variety in economic theory</b>	<b>11</b>
2.1	The origins of monopolistic competition . . . . .	11
2.2	Product variety and consumer welfare . . . . .	12
2.3	The Dixit-Stiglitz approach to monopolistic competition . . . . .	14
2.4	Other theories involving product variety . . . . .	18
	Notes . . . . .	19
<b>3</b>	<b>Empirical trends</b>	<b>21</b>
3.1	Trade, income convergence, and scale . . . . .	21
3.2	Education and research . . . . .	23
3.3	Product variety . . . . .	29
3.4	Summary . . . . .	33
	Notes . . . . .	35
<b>4</b>	<b>Knowledge spillovers and population growth: Endogenous growth theory 1995-2003</b>	<b>37</b>
4.1	Introduction . . . . .	37
4.2	Production . . . . .	40
4.3	Research . . . . .	42
4.4	Households . . . . .	47
4.5	Aggregation . . . . .	50
4.6	Endogenous growth . . . . .	51
4.7	Semi-endogenous growth . . . . .	54
4.8	Neo-endogenous growth . . . . .	55
4.8.1	Comparative static analysis . . . . .	56
4.8.2	Dynamic analysis . . . . .	57
4.8.3	Network-confined spillovers . . . . .	58
4.8.4	Type (iii) spillovers . . . . .	59
4.9	Conclusion . . . . .	59
4.A	Type (iv) spillovers . . . . .	60
	Notes . . . . .	61

<b>5</b>	<b>A measure of product variety</b>	<b>63</b>
5.1	Introduction . . . . .	63
5.2	Entropy and product variety . . . . .	64
5.3	Discussion . . . . .	67
5.4	Conclusion . . . . .	68
	Notes . . . . .	69
<b>6</b>	<b>Indices without baskets</b>	<b>71</b>
6.1	Introduction . . . . .	71
6.2	Variety robust indices . . . . .	73
6.3	Simulation results . . . . .	76
6.4	Applications to economic theory . . . . .	79
6.4.1	Productivity measurement . . . . .	79
6.4.2	CES functions . . . . .	80
6.4.3	Cost of living indices . . . . .	80
6.5	Conclusion . . . . .	82
6.A	Multiplicative decomposition of additive indices . . . . .	82
6.B	Application to CRESH functions . . . . .	83
6.C	Comment on impossibility theorems . . . . .	84
6.D	Feenstra's variety index . . . . .	85
	Notes . . . . .	86
<b>7</b>	<b>Two models of economic growth</b>	<b>87</b>
7.1	Introduction . . . . .	87
7.2	Schooling inequality and growth . . . . .	87
7.2.1	Introduction . . . . .	87
7.2.2	The model . . . . .	89
7.2.3	Steady state . . . . .	92
7.2.4	Schooling inequality . . . . .	94
7.2.5	Summary . . . . .	99
7.3	Imperfect improvements . . . . .	99
7.3.1	Introduction . . . . .	99
7.3.2	The model . . . . .	102
7.3.3	Discussion . . . . .	104
7.4	Conclusion . . . . .	106
7.A	Endogenous skill growth . . . . .	107
7.B	Endogenous growth with imperfect improvements . . . . .	108
7.C	Derivation of $V$ . . . . .	108
	Notes . . . . .	109
<b>8</b>	<b>Economic integration</b>	<b>111</b>
8.1	Introduction . . . . .	111
8.2	The model . . . . .	114
8.3	Aggregate welfare . . . . .	117
8.4	Income inequality . . . . .	120
8.5	Conclusion . . . . .	126
8.A	Stability of the model . . . . .	127
	Notes . . . . .	129

<b>9</b>	<b>Idiosyncratic firm performance</b>	<b>131</b>
9.1	Introduction . . . . .	131
9.2	Discrete model and asymptotic results . . . . .	132
9.3	Small sample results . . . . .	135
9.4	Continuous model . . . . .	136
9.5	Conclusion . . . . .	139
9.A	Simulation results . . . . .	141
9.B	Mathematical appendix . . . . .	146
9.B.1	Theorems . . . . .	146
9.B.2	Derivation of $C$ in continuous time . . . . .	148
9.B.3	Derivation of $N$ in continuous time . . . . .	149
	Notes . . . . .	150
<b>10</b>	<b>Empirical analysis</b>	<b>151</b>
10.1	Introduction . . . . .	151
10.2	Product variety and country size . . . . .	152
10.3	The curse of classification? . . . . .	158
10.4	Longitudinal determinants of product variety . . . . .	162
10.5	Idiosyncrasy and effective variety . . . . .	172
10.6	Conclusion . . . . .	175
10.A	Additional figures and tables . . . . .	176
	Notes . . . . .	180
<b>11</b>	<b>Conclusion</b>	<b>183</b>
11.1	Summary . . . . .	183
11.2	Further research . . . . .	185
11.3	The Great Convergence . . . . .	186
	Notes . . . . .	189
<b>A</b>	<b>Data on product variety</b>	<b>191</b>
	<b>Bibliography</b>	<b>195</b>
	<b>Index</b>	<b>205</b>
	<b>Samenvatting (Summary in dutch)</b>	<b>209</b>
	<b>Curriculum Vitae</b>	<b>211</b>



# List of Tables

3.1	Relative sectoral size during the British Industrial Revolution . . . .	30
3.2	Trends in macroeconomic variables . . . . .	34
4.1	Parameterization of production . . . . .	41
4.2	Parameterization of research . . . . .	46
4.3	Four special cases of growth models . . . . .	55
8.1	Eigenvalues and eigenvectors for linearized system . . . . .	128
9.1	Simulation setup . . . . .	135
10.1	Regression product variety and population (cross-section) . . . . .	154
10.2	Regression estimated product variety and population (cross-section)	160
10.3	Approximation of production data . . . . .	163
10.4	Trend export variety SITC 2 . . . . .	165
10.5	Trend export variety per capita SITC 2 . . . . .	167
10.6	Trend export variety per highly educated person SITC 2 . . . . .	169
10.7	Summary trends in number of product groups . . . . .	170
10.8	t-test of $\hat{g}_X - \hat{H} > 0$ . . . . .	170
10.9	Trend effective variety per product group . . . . .	174
10.10	Trend ratio export variety to years of higher schooling SITC 2 . . . .	177
10.11	Trend export variety SITC 3 . . . . .	178
10.12	Trend export variety HS 96 . . . . .	179
A.1	ISOALPHA 3 country codes . . . . .	192
A.2	Data coverage SITC rev. 2 and 3 . . . . .	193
A.3	Data coverage HS 96 . . . . .	194





# List of Figures

3.1	Exports relative to GDP (1960-2000) . . . . .	22
3.2	Imports relative to GDP (1960-2000) . . . . .	22
3.3	GDP per capita and population (1960-2000) . . . . .	23
3.4	GDP per capita and population (OECD aggregate; 1970-2000) . . . . .	24
3.5	Years of higher schooling per capita (1960-2000) . . . . .	25
3.6	Tertiary education completed above 25 (1960-2000) . . . . .	25
3.7	Primary, secondary, and tertiary education completed (1960-2000) . . . . .	26
3.8	Gross expenditure on R&D relative to GDP (1981-1999) . . . . .	26
3.9	Number of researchers per worker (1981-1999) . . . . .	27
3.10	Patent flows (1981-1999) . . . . .	27
3.11	Number of patents per researcher (1981-1999) . . . . .	28
3.12	Number of traded product groups by country (1961-2001) . . . . .	31
3.13	Number and value of traded product groups (1961-2001) . . . . .	32
3.14	Rescaled number of traded product groups (1961-2001) . . . . .	32
5.1	Entropy and product classification . . . . .	65
6.1	Simulation results quantity indices with increasing variety . . . . .	77
6.2	Simulation results quantity indices with decreasing variety . . . . .	77
6.3	Simulation results price indices with increasing variety . . . . .	78
6.4	Simulation results price indices with decreasing variety . . . . .	78
7.1	Density plots for skill distributions . . . . .	90
7.2	Inequality and skills . . . . .	95
7.3	Income and skill change . . . . .	95
7.4	Job choice by worker $L_x$ (constant wages) . . . . .	96
7.5	Adjustment of wage ratio to steady state . . . . .	97
7.6	Job choice by worker $L_x$ : rising inequality . . . . .	97
7.7	Job choice by worker $L_x$ : declining inequality . . . . .	98
7.8	Product variety in the steady state . . . . .	99
8.1	Per capita consumption – small country . . . . .	119
8.2	Number of types consumed – small country . . . . .	120
8.3	Scale of production . . . . .	120
8.4	Per capita consumption – large country . . . . .	121
8.5	Number of types consumed – large country . . . . .	121
8.6	$Y_{pc}$ (solid) and $C_{pc}$ (dashed) by income group . . . . .	122
8.7	Net asset income . . . . .	123
8.8	Wage inequality . . . . .	124
8.9	$Y_{pc}$ (solid) and $C_{pc}$ (dashed) by income group; shock in $H/L$ . . . . .	125
8.10	Wage inequality; shock in $H/L$ . . . . .	125

8.11	Scale of production; shock in $H/L$ . . . . .	126
9.1	Sample variance of productivity . . . . .	141
9.2	Aggregate output . . . . .	142
9.3	Product variety . . . . .	143
9.4	Consumption per capita . . . . .	144
9.5	Consumption value per unit of production . . . . .	145
10.1	Theoretical overview variety and population . . . . .	153
10.2	Product groups and population for US states . . . . .	156
10.3	Establishments and population for US states . . . . .	156
10.4	Product groups and population for European countries . . . . .	157
10.5	Establishments and population for European countries . . . . .	157
10.6	Estimated product variety for US states . . . . .	159
10.7	Estimated product variety with clustering . . . . .	161
10.8	Estimated product variety with clustering and population size . . . . .	161
10.9	Number of product groups per highly educated person . . . . .	168
10.10	Effective variety per product group . . . . .	172
10.11	Product groups and population (SITC 3) . . . . .	176
10.12	Product groups and population (HS 96) . . . . .	176

## Conventions and commonly used definitions

---

### Conventions

---

$\dot{x} \equiv \frac{dx}{dt}$	Derivative of $x$ with respect to time
$\hat{x} \equiv \frac{\dot{x}}{x} = \frac{d \ln x}{dt}$	Growth rate of $x$
$g_x = \hat{x}$	Constant (steady state) growth rate of $x$
$\varepsilon_{yx} \equiv \frac{d \ln y}{d \ln x}$	Elasticity of $y$ with respect to $x$
$E(x)$	Expected value of $x$
$\mu(x), m(x)$	Asymptotic mean of $x$ , sample mean of $x$
$\sigma^2(x), s^2(x)$	Asymptotic variance of $x$ , sample variance of $x$
Upper-case roman	Aggregate variables; indices; sets
Lower-case roman	Disaggregate variables; non-exponent parameters
Upper-case greek	Expressions used in order to simplify algebra
Lower-case greek	Parameters (mostly in exponents)

### Commonly used definitions

---

$A$	Stock of ideas; mass of product types
$C, c$	(Aggregate) consumption; cost of living index
$E, e$	(Aggregate) expenditure
$G, g$	Set of product groups, number of product groups
$H, h$	(Aggregate) human capital; entropy
$i, j$	Index
$k$	Skills
$K$	Number of networks
$L, l$	(Aggregate) labor
$N$	Set of product types; variety index (also $V$ )
$n$	Number or mass of product types
$P, p$	(Aggregate) population; price index, price
$Q$	Aggregate quantity index; set of ideas
$q$	Quality (occasionally $\theta$ )
$s$	Average skill level
$t$	Time
$U$	Net present value of instantaneous utility
$u$	Instantaneous utility
$V$	Variety index (also $N$ )
$w$	Wage rate; weight, budget share
$X$	'Average' quantity index
$x$	Quantity of (intermediate) output
$Y, y$	(Aggregate) income; (aggregate) quantity of final output
$1/(1 - \gamma)$	Elasticity of substitution
$\varepsilon$	Parameter for knowledge spillovers
$\Theta, \theta$	Quality index, quality (usually $q$ )
$\Pi, \pi$	(Aggregate) profit
$\rho$	Subjective discount rate
$\sigma$	Inequality of skill distribution
$\omega$	Wage rate skilled labor

---



# Introduction

## 1.1 Variety is the spice of life

Spices were in great demand in medieval Europe. Their popularity had to do with the inferior techniques for the preservation of food in those days. In order to save grain, a large part of the stock of cattle was slaughtered in the autumn; the meat was stored after being smoked, corned, or spiced (Landes 1998, p. 132-3). These means of preservation did not work very well and the only way to make a winter's meal bearable was to add spices to it. Besides being in great demand, spices were also in short supply because they came from far (India, the Moluccas) and the means of transportation were slow. Initially, Europeans bought their spices in the ports of Egypt and other countries bordering the Middle-East. The goods arrived there after they had passed through the hands of several middle-men, each demanding his share. In the 16th century, after Vasco da Gama had rounded Cape Good Hope, the Portuguese started to buy (and rob) directly from the source. The Dutch and English took over their place in the 17th century.

Before the 19th century, trade was largely restricted to non-competitive goods (O'Rourke and Williamson 2000, p. 5). These goods were imported because they could not be produced domestically, either for geographical reasons (spices, precious metals) or because the production process was unknown (porcelain). European import from Asia thus mainly consisted of spices, coffee, tea, silk, and sugar, while silver, linens, and woolens were exchanged in return. The primary reason for trade had been the expansion of the variety of consumer goods, or, formulated more strongly, trade enabled product variety.

O'Rourke and Williamson consider the period between 1400 and 1800 to be the first of three eras in global trade. After the era of trade in non-competing goods, a second era (1800-1900) began, characterized by trade in 'basic' competing goods. The third era (1900-present) displayed trade in highly differentiated manufactured goods.

At the end of the first era, the idea that the location of production need not be geographically fixed gained ground in Europe. Consider hard porcelain, for example. Although the Portuguese had already imported hard porcelain from China in the 1520s, it was only in 1715 that Böttcher, located in Dresden, reinvented hard porcelain (Marryat 1857, pp. 185-7, 241-3). This reinvention was the result of well-funded and focussed research (as well as pure luck!). Imitation gradually became a substitute for trade.

The end of the Napoleonic Wars, around 1820, marked the beginning of the second era of trade in basic competing goods. The abolition of the Corn Laws and various international treaties created a more favorable environment for trade. From 1870 onwards, trade was further stimulated by a dramatic fall in transport costs and by the spread of the gold standard (Estevadeordal, Frantz, and Taylor 2003). In this second era, the cost of production began to play an important role as a motive for trade. Were the traded commodities in the 17th century mainly exotic luxury goods, in the 18th century large amounts of ordinary goods were traded, such as textiles and wheat. The decline in transport costs and the spread of imitation reduced the importance of the location of production, and made trade more sensitive to a country's relative endowment of nontradable inputs. In this Heckscher-Ohlin era, trade not only served the purpose of expanding the product range but it also led to foreign competition for domestically produced goods, and consequently to lower prices for these goods.

The third era of global trade is that of trade in highly differentiated products. O'Rourke and Williamson suggest that this era covers the 20th century, but considering the isolationism in Europe and the two World Wars of the first half of the 20th century, it seems more appropriate to let this era begin in the 1950s when trade liberalization became popular again.<sup>1</sup> International trade in the third era became more and more similar to intra-national trade – at least for the economically advanced capitalistic countries. This group of countries, broadly the members of the Organization for Economic Cooperation and Development (OECD), formed an integrated economy in which per capita incomes are roughly identical.<sup>2</sup> It is this era and these countries that have been studied in this dissertation, although some conclusions might be applicable to the earlier eras as well.

Section 1.2 of this chapter introduces the main assumption and the two main propositions of this dissertation. Besides an assessment of the propositions, a number of results can be found in this dissertation that are useful in their own right. The most important of these results are mentioned in section 1.3, which provides an overview of the remaining chapters of the dissertation.

## 1.2 One assumption and two propositions

The two main propositions of the dissertation pertain to two different mechanisms that affect product variety. Their theoretical foundations and empirical plausibility are evaluated in the chapters that are to come. Before turning to these propositions, however, the main assumption of the dissertation is introduced.

### 1.2.1 New goods and the variety of consumer products

The objectives of consumption can be divided into four categories. First, consumer goods – when I use the terms 'goods' or 'products' I implicitly refer to services as well – may improve or maintain the health of a person. Trivially, food and clothing are consumed partly for this reason. Another example is health care. Second, the use of some goods may free a substantial amount of time (Becker 1965). Examples of goods consumed for this reason include spectacles, washing machines, cars, cell phones, and light bulbs.<sup>3</sup> A third reason, of course, is recreation or entertainment, which is provided by restaurants, cd's, televisions, and chatting on the internet, to name a few examples. This kind of consumption can be labeled 'hedonic' consumption. A fourth reason is 'conspicuous consumption'.<sup>4</sup> People consume

goods – like clothing and cars (again) – to influence their social status. (I do not consider expenditure with altruistic objectives to be consumption.)

As is immediately clear from the examples above, a good can have multiple purposes, but it usually is not difficult to discern a dominant reason why people consume a particular type of good. A car is mainly bought to save time, and the primary function of a television is entertainment.

When a person wishes to improve his situation in one or more of the 'dimensions' of consumption mentioned above, just consuming more of the same types of goods is usually not very effective. Two washing machines do not make cleaner clothes than one does. Instead, it is more effective to consume types of goods that were not consumed before. These new types may replace the previously consumed types, or not.

In many cases, the invention of a new good did not lead to the extinction of its predecessors, even when the new good is superior in many respects. The invention of the light bulb does not imply that candles are not used anymore and people still build sailing yachts although motorized ships and airplanes are much more efficient forms of transportation. The widespread adoption of computers did not make books and other printed matter obsolete; it did not even lead to the extinction of calculators!

If the invention and production of a new good does not lead to the disappearance of one or more existing goods, then the number of products in the economy must increase. In principle, there are two kinds of reasons why new goods may not be perfect substitutes, or rather superior substitutes, for existing goods. First, the new good can be seen as an imperfect substitute by the individual consumer. Such products can be thought of as being suited for some activity that could not be properly performed by existing products. A straightforward example is the invention of a medicine for a disease for which no treatment had been available previously. New goods may perform different functions than those performed by existing goods.

Second, a new product can be seen as a superior substitute for some existing product by some consumers, while other consumers prefer the existing product. The introduction of a new good may increase the number of product types in the economy because of heterogeneity of preferences.

For the reasons discussed above, the invention of new goods simultaneously leads to a higher standard of living and a larger number of products. Two arguments can be made against this statement. First, new goods do not necessarily raise the standard of living because their consumption or production might have negative effects that are ignored by the consumer and the producer.<sup>5</sup> The emission of greenhouse gasses due to the use of energy is an important example of such a negative 'externality'. Similarly, it is not easy to defend the position that new types of weapons generally contribute to social welfare. This is an important argument but it is not likely to reverse the relation between new goods and the standard of living as externalities can often be curtailed by government policies and other forms of collective action.

Second, one may argue that a new type of product causes the extinction of one or more older types and therefore might even lead to a decrease in the number of products being consumed. This is particularly relevant when the novelty of a product is appreciated by consumers, a common phenomenon with hedonic and conspicuous types of consumption. Obviously, this argument cannot hold in general as this would imply that the number of products consumed a thousand years ago would be larger than the current number of products. This brings me to the main assumption of the dissertation.



**Assumption 1** *The invention of new types of products raises both the standard of living and the number of product types that are being produced.*

A distinction between consumer goods and intermediate inputs is intentionally avoided here. Although the determinants of the variety of consumer goods are the primary concern of this dissertation, most of the analysis presented in the remaining chapters is also applicable to intermediate products. The variety of intermediates may raise the quality or lower the production costs of final goods, thereby indirectly affecting the standard of living.

## 1.2.2 Product variety and the size of the economy

While it might be quite obvious that new goods raise the standard of living, it is less obvious what causes a new good to be invented and produced. Putting aside pure chance, an invention will only take place if the would-be inventors have inspiration and the time and tools to convert an idea into a working prototype. Whether a new invention will actually be produced on a large scale, depends purely on its profitability.

Inspiration is affected by previous inventions. One possibility is that the best inventions – the wheel – are made first. The inspiration needed for the invention of new products would then decline over time. Jones labeled this effect ‘fishing-out’ (Jones 1995a). A similar argument is that the probability of duplicate ideas rises over time. Another possibility is that past inventions inspire current inventors by providing them with more knowledge. Wilbur and Orville Wright, for example, benefitted from their experience as bicycle manufacturers when constructing the first airplane that was capable of sustained, controlled powered flight. By this line of argument, the inspiration for new types of products must rise over time. This effect is labeled ‘standing on the shoulders of giants’ after a quote from Isaac Newton.

(It should be noted that the ‘fishing out’ and ‘standing on shoulders’ effects are not mutually exclusive phenomena. A major discovery of the ‘standing on shoulders’ type – a new ‘fishing-pond’ – may very well trigger a sequence of smaller inventions of the ‘fishing out’ type. In particular, a new general purpose technologies (GPTs) bring along opportunities for complementary inventions (Bresnahan and Trajtenberg 1995).)

Besides inspiration, perspiration is the other condition for invention. Without a considerable amount of effort, a new idea will not become a new product. Thus, invention is a costly process and has to be funded in some way. Sometimes researchers fund themselves, like the Wright brothers, but nowadays research is usually funded by companies or governments.

Whether a research project is funded or not, depends on the rewards the invention is expected to generate relative to the expected costs of the project. There are two factors that determine the expected costs of a project: the expected time of development, and the prices of inputs needed for research (usually the wages of researchers). The first factor, the development time, depends on the number and productivity of researchers hired, while the last factor, the wage rate of researchers, depends on the scarcity of highly educated workers. Research is therefore likely to be less costly if the workforce is better educated.

The expected rewards are influenced by three factors: the probability that an invention occurs, the extent to which the invention improves on existing inventions (the functionality of the invention), and the scale at which the invention can be applied. Additionally, the extent to which the benefits of invention can be

appropriated by the owner of the invention is important when research is funded by companies.<sup>6</sup>

The first two factors, the probability of success and the functionality of the invention, are related to the quality of the initial idea and, hence, to the stock of previous inventions. The third factor, the scale of application, is bounded by the size of the market for the newly invented product. The size of the market, in turn, is largely determined by the number of consumers in the economy and the stock of previously invented products.

For any given stock of inventions, an economy with a larger and better educated workforce will generate a larger number of new goods than an economy with a smaller and less well educated workforce. This suggests that the number of products can grow in three ways. First, the number of products can grow gradually because population growth ensures the expansion of markets. Second, it can grow rapidly when two or more economies are merging. Third, advances in education may stimulate growth as they raise the productive capacity of the workforce.<sup>7</sup> This line of reasoning forms the basis for the first proposition of the dissertation.

**Proposition 1** *Population growth, economic integration, and a better educated workforce all have contributed to a larger variety of product types in OECD countries.*

### 1.2.3 Product variety and idiosyncratic firm performance

The relation between product variety and the standard of living does not have to be positive. In particular, when the performance of a firm or industry in terms of the costs and quality of its products is prone to persistent idiosyncratic shocks, a selection mechanism similar to Fisher's fundamental theorem of natural selection (Fisher 1958 [1930]) comes into effect.

If the effects of shocks last for a long time, the difference in the performance of firms will grow over time. In order to see how this works, suppose that all firms start out with the same level of performance and administer some random and firm-specific shocks for some time. At the end of this experiment, a few firms might have benefitted from most of the shocks they got, while others saw their performance decline with almost each shock. Many firms will have had a mixed experience. This simple experiment suggests that as time proceeds, the differences between 'lucky' and 'unlucky' firms will widen.

The heterogeneity in firm performance caused by idiosyncratic shocks has implications for the quantity sold by individual producers. Consumers will tend to buy more from the better performing firms, while they will tend to consume less of the products produced by the badly performing firms. This will affect the standard of living in two ways. First, consumers are able to buy a larger quantity of goods, assuming that firms differ with respect to their technical efficiency. Naturally, this will raise the standard of living. Second, it has become more costly for consumers to buy certain types of goods, which implies a loss in the variety of goods as it is experienced by consumers. This loss in variety has a negative effect on the standard of living.

In order to clarify why an increase in the heterogeneity of costs causes a decline in variety, let me elaborate on exactly what I mean with variety here. The most intuitive measure of product variety is the number of product types. A notable disadvantage of this measure is that it implicitly assumes that all types are weighted equally by consumers. To get a grasp of why weights matter for the way a consumer

perceives variety, compare the following two situations. In the first situation, a consumer assigns equal weights (e.g. budget shares) to all available product types. If the consumer would be asked to abstain from consuming one type of good, she would be indifferent about which type to choose. In the second situation, she assigns large weights to some types and small weights to other types. Now she will choose a type with a small weight as this choice has the smallest effect on her utility. In both situations the number of product types is reduced by one, but the consumer clearly suffers the greatest loss in the first case. Hence, variety becomes more important to a consumer when the weights assigned to individual product types become more equal. I will label product variety as it is experienced by consumers – thus taking into account how consumers weigh product types – ‘effective variety’.

A fundamental characteristic of effective variety is that it might change while the number of product types remains constant. The weights consumers attach to product types might change due to, for example, changes in prices. In particular, a product may become so expensive that the share of a consumer’s budget it receives becomes negligible. Such an event does not affect the number of types, although effectively the consumer uses one type less. A divergence in the productivity of firms leads to a concentration of expenditure on a subset of types, thus causing a reduction in effective variety even though the number of types does not change.

This principle is not restricted to heterogeneity in technical efficiency: it also applies when firms experience shocks in the quality of their products. When firms differ with respect to the quality of their products, consumers will buy more of the goods with a higher quality. This raises the standard of living. On the other hand, the quality of some product types – those from ‘unlucky’ firms – will be inferior.

By Fisher’s theorem, rising standards of living may coincide with a declining effective variety of consumer goods when the performance of firms or industries becomes more heterogeneous over time.

**Proposition 2** *The differential performance of industries in OECD countries has had a negative effect on the effective variety of products.*

The scale effect of proposition 1 and the variation and selection effect of proposition 2 suggest opposite relations between product variety and the standard of living. The main purpose of this dissertation is to assess the theoretical and empirical plausibility of the two propositions and to provide evidence on which of the two mechanisms is dominant.

### 1.3 Research outline

The perspective on the relation between product variety and economic growth reflected by propositions 1 and 2 differs from the perspective commonly adopted in the theoretical literature on economic growth. In this strand of literature, product variety is assumed to be determined by two factors: population growth and intertemporal knowledge spillovers. Population size is important as the costs of R&D are more easily recovered when a newly invented product can be sold to a larger number of people. This mechanism was first proposed by Judd (1985b). Later, Romer (1990) augmented Judd’s model with intertemporal knowledge spillovers (i.e. the effects of previous research outcomes on the productivity of researchers). Intertemporal knowledge spillovers reduce the amount of labor needed for the invention a new type of good. For example, the invention of the computer has made

(at least some) scientists more productive. If these spillovers are strong enough like in Romer's model, then population growth is no longer necessary for sustained growth in product variety. Jones (1995a) has shown that also weak spillovers affect the growth rate of product variety, but only when population growth is positive.

In a nutshell, the research presented in this thesis deviates from the existing literature in three respects. First, intertemporal knowledge spillovers are considered to be relatively insignificant (evidence that supports this assumption is presented at the end of the thesis). Second, it is emphasized here that in addition to population growth also advances in education and ongoing economic integration raise the scale of the 'world-economy'. Third, attention is paid to the causes and consequences of asymmetry in the sales of product types. Idiosyncratic shocks to the performance of firms provide one reason why not all product types are sold in equal amounts to consumers. If not every type of good is consumed by the same amount, this will affect the way product variety is experienced by consumers. The existing literature does not consider this phenomenon.

The three chapters subsequent to the current chapter are still of an introductory nature. Chapter 2 provides an overview of the theoretical literature on product variety. A large part of this chapter is devoted to the history of the theory of monopolistic competition. Monopolistic competition is closely linked to product variety because entrepreneurs like to sell products that differ from their competitor's products. The theory of monopolistic competition can therefore be considered one of the cornerstones of the modern theory of economic growth. The survey starts with the early works by Sraffa and Hotelling and, via the Dixit-Stiglitz revolution, ends with the growth models by Romer (1990) and Jones (1995a).

Some important empirical trends for the OECD are presented in chapter 3. The perspective taken in this chapter differs somewhat from the perspective that is customary in the empirical literature on economic growth in that it considers the members of the OECD to collectively form a single integrated economy. The trends are subdivided in three themes. The first theme comprises trade, income, and population size; education and research form a second theme, and the third theme is product variety. The empirical trends set the stage for theoretical models presented later on and, therefore, they will be referred to frequently.

Chapter 4 treats the literature on economic growth from 1995 onwards.<sup>8</sup> The chapter starts by introducing the 'Jones critique', which states that the long-term rate of economic growth should not depend on population size. A controversial feature of the model that Jones proposed is that economic growth requires population growth because knowledge spillovers are not strong enough. From 1998 onwards a number of models have been suggested that complied with the Jones critique, while still allowing for knowledge spillovers that are sufficiently strong to ensure economic growth when population size is constant.

Because this literature evolves largely around technicalities, it is 'reviewed' by analyzing a generalized growth model. This approach not only has the advantage that it facilitates the comparison of various contributions to the literature, it also yields some new insights in how these models work. Several standard results derived in chapter 4 will be referred to in later chapters in order to avoid repetition.

Chapters 5 and 6 are primarily concerned with measurement issues related to product variety. Chapter 5 introduces an indicator of effective product variety in the form of the antilog of Shannon's entropy. Such an indicator is useful if not all product types can be considered to be equally important. The proposed variety index also plays a role in chapter 6. In this chapter, a new set of indices for prices and quantities is presented. These price and quantity indices have the advantage

over established indices that changes in product variety do not cause them to be biased. It is shown that the indices satisfy a set of important axioms, and that they can conveniently be combined with economic theory. In particular, the indices can be used to estimate the welfare effects stemming from changes in product variety. The proposed method for doing so is more general than the method suggested by Feenstra (Feenstra, Markusen, and Zeile 1992, Feenstra 1994, Feenstra and Markusen 1994, Feenstra and Kee 2004).

Chapter 7 contains two alternatives to the models discussed in chapter 4. The first model shows how schooling may affect the number of product types in an economy. Growth in the average level of schooling is demonstrated to stimulate the invention of new products in the same way as population growth – namely by increasing the scale of the economy. Besides the average level of schooling, also the distribution of schooling is shown to be relevant. Under certain conditions, a decline in schooling inequality can lead to a rise in the proportion of researchers in the population. A different explanation is provided by Jones (2002), who suggests that the rise of research is due to growth in the average level of education.

The second model of chapter 7 covers the spectrum between variety expansion and quality improvements in the sense that each new product type has a quality that is higher than that of the existing types. The model is the first of this dissertation in which the number of product types exceeds effective variety.

Chapter 8 treats the effects of economic integration on product variety and the distribution of income. The model that is used in this chapter combines economic growth driven by the invention of new types of products with fixed costs of production. These fixed costs of production give rise to economies of scale at the firm level, which provide an incentive for the geographical concentration of production and thus for international trade. Prior to integration countries imitate each other's products; after integration the incentive to do so disappears.

Chapter 9 demonstrates how idiosyncratic firm performance may reduce the effective variety of consumer goods while raising the standard of living. The chapter starts by illustrating the basic mechanism using a model in discrete time. Results are presented for both a continuum and a discrete number of firms. At the end of the chapter idiosyncratic productivity shocks are introduced into a simple growth model in continuous time.

The models of chapters 7 and 8 point out how advances in education and economic integration stimulate the growth of product variety through their effect on the scale of the economy. For this reason these chapters are of particular interest for proposition 1. Chapter 9 provides the theoretical basis for proposition 2.

Chapter 10 confronts the theoretical considerations outlined above with data on product variety. Both a cross-sectional and a time-series analysis are conducted. The cross-sectional analysis suggests that the scale of production is an important determinant of the number of produced product types in small countries, while produced variety in large countries seems to be constrained by the worldwide stock of inventions. The time-series analysis shows that within countries, product variety can largely be explained by the stock of human capital. Furthermore, the ratio of effective variety to the number of product types is found to be declining in most countries, which points to the presence of idiosyncratic industry performance.

Chapter 11 summarizes the conclusions of the other chapters and returns to the two propositions stated above.

## Notes

1. One might also add a fourth era for the trade in services, which probably took off in the 1990s.
2. See Ben-David (1993) for a study of the consequences of trade liberalization.
3. The invention of spectacles in medieval Europe enabled craftsmen to continue working well after the age of forty (Landes 1998, p. 46). Nordhaus (1997) shows that the price of lighting has dropped enormously the last two hundred years. This made activities feasible that were previously too expensive to be performed in the dark.
4. The classical references on conspicuous consumption include those to Rae (1905 [1834]), Veblen (1921), Duesenberry (1949), and Leibenstein (1950). For a survey of the literature, see Reinstaller and Sanditov (2003).
5. Schwartz (2004), for example, argues that having more options may actually cause a deterioration of a person's well-being.
6. The appropriation of the benefits of invention strongly depends on the ability of competitors to imitate the invention, or, alternatively, on their ability to discover a superior design. Competition might come from newly started firms led by entrepreneur-inventors (like in the biotech industry) or it might come from incumbent firms that typically rely on a formal R&D-department for their 'supply' of new products. The first type of competition has been labeled 'Schumpeter Mark I' after Schumpeter's early view on market dynamics (Schumpeter 1934). The second type has been labeled 'Schumpeter Mark II' because it corresponds with a more recent perspective of Schumpeter on the dynamics of competition (Schumpeter 1943).
7. A more controversial explanation of economic growth is that technological change has such a large impact on the productivity of researchers that economic growth becomes a self-reinforcing (autocatalytic) process (Romer 1990).
8. Surveys of the literature up to 1995 are provided in Barro and Sala-i-Martin (2004), Jones (1998), and Aghion and Howitt (1998).



## Product variety in economic theory

This chapter starts with the origins of monopolistic competition and its early developments. The theory of monopolistic competition has proven to be an essential ingredient of theories with an endogenous number of products (Klundert and Smulders 2001). After this exposition the various (early) perspectives on the relation between product variety and consumer welfare are discussed in section 2.2. Section 2.3 of this chapter contains a description of the Dixit-Stiglitz framework and its application in theories of international trade and economic growth. The chapter concludes with a brief survey of some theories involving product variety that are unrelated to monopolistic competition.

### 2.1 The origins of monopolistic competition

In the late 1920s economists began to feel uncomfortable with the established theories of competition. The main reason for their discomfort was that in all of these theories any seller could capture the entire market by setting a price just below that of its competitor(s). The fact that identical or similar goods were simultaneously traded at different prices could not be explained by the theories of monopoly, oligopoly, or perfect competition. Sraffa (1926) was probably the first author who recognized that heterogeneity in preferences leads to “a willingness (..) to pay (..) something extra in order to obtain the goods from a particular firm rather than from another” (p. 545). Not only did Sraffa emphasize that competition is rarely perfect, he also put forward that monopoly is rarely perfect as well.

... the majority of circumstances which affect the strength of a monopolist (..) exercise their influence essentially by affecting the elasticity of demand for the monopolized goods. Whatever the causes may be, this is the only decisive factor in estimating the degree of independence which a monopolist has in fixing prices: the less elastic the demand for his product, the greater is his hold on his market. (Sraffa 1926, p. 545)

On the same page Sraffa explicitly states that a higher elasticity of demand implies a higher degree of competition and that an infinite elasticity of demand corresponds to perfect competition.



In his article 'Stability in competition' Hotelling (1929) presents a formal theory of 'quasi-monopoly' in which each firm has a cost advantage over other firms for a group of customers that are located near the firm. In Hotelling's model, customers do not only take into account the price of the product but also the location of the seller because buying from a more distant firm brings along higher transportation costs. The geographical dispersion of customers gives sellers power over a part of the market such that a small change in price does not cause all the customers of the slightly more expensive firm to defect to the cheaper firm. For some buyers, the price difference might not be large enough to compensate for the rise in transportation costs associated with buying from the other seller. Hotelling's model of competition thus differed from existing models of competition in that it allows for sustained differences in the price of identical goods.

Hotelling's model of spatial competition fitted into the tradition of using the geographical location of buyers and sellers as an analogy of heterogeneity in preferences and products, respectively. The distance between a buyer and a seller can be considered the distance between the product being on sale and the 'ideal product' of the buyer. A rise in the number of sellers then enables a better fit between products and preferences in a similar fashion as a greater number of sellers reduces the average costs of transportation in the spatial model. Lancaster (1966, 1975, 1979) and later Salop (1979) began to adopt this analogy explicitly. Lancaster, in particular, generalized Hotelling's model to more than one dimension. In his model products have several characteristics in stead of just one (for example the distance between buyer and seller).

In 1933 two influential books were published: 'The theory of monopolistic competition' by Chamberlin (1962 [1933]), which was based on his thesis from 1927, and 'Economics of imperfect competition' by Robinson (1950 [1933]). Although both authors were concerned with market structures that are in between the stereotypes of 'pure monopoly' and 'perfect competition', they emphasized different aspects. Robinson was primarily concerned with market power arising from 'non-infinite elasticities of demand', or in other words imperfect competition.<sup>1</sup> Chamberlin was more concerned with how differences between products affects the competition between firms.<sup>2</sup> According to Chamberlin competition and monopoly are not mutually exclusive. A monopolist being the only seller of a particular product cannot wield unlimited power. A price set too high will cause its customers to spend their money on other products, similar to those of the monopolist or not. Chamberlin labeled the competition between monopolists *monopolistic competition*.

## 2.2 Product variety and consumer welfare

The discovery of monopolistic competition brought with it the peculiar phenomenon of 'excess capacity' in competitive equilibrium. The term 'excess capacity' refers here to production without full exploitation of economies of scale so that average costs are higher than minimum average costs.<sup>3</sup> Under perfect competition (pure competition in Chamberlin's words) firms producing less than the optimal amount will disappear due to the entry of new firms. The higher the minimum average cost level relative to the size of the market, the smaller the number of firms that can be sustained in it. In the limiting case where economies of scale are practically unbounded, competition would be unstable and a single firm would capture the entire market – with zero profits because of the threat of entry.

Chamberlin showed in section 5.5 of his book that in a situation of monopolistic competition, the equilibrium level of production might differ from that of the level associated with minimum average costs. Chamberlin's analysis is essentially graphical but his line of reasoning can be summarized as follows. Each firm, being a monopolist for its own product, sets a price higher than average costs. The profits obtained by the firm will induce the 'entry' of new firms, each of them producing their own type of product. The consequence of the increase in the number of firms is that each firm sees the demand for its product decrease. The reduction in demand causes the firm to produce less and thus average costs will be higher than before. The incentive for entry will disappear as soon as the number of firms is such that profits are zero.

Under perfect competition entry reduces the price and the incumbent firms have no option other than to follow the market price while keeping average costs to a minimum. Under monopolistic competition firms respond to entry by reducing the level of production. Whereas entry under perfect competition weeds out inefficient firms, entry under monopolistic competition has the opposite effect of raising average costs.

Chamberlin's discussion of 'excess capacity' in competitive equilibrium did not pass unnoticed. In 1935 Kaldor dedicated an article to this topic, in which he pointed out that an increase in product variety might be detrimental to consumer welfare due to the loss of scale economies. He clearly is aware of the trade-off between 'quantity and diversity', but he is dismissive of the argument that the level of 'excess capacity' with free entry is optimal as it is the result of consumer choices.

This line of reasoning would only be permissible if consumers were actually confronted with the choice of having *either* a smaller range of commodities at lower prices *or* a larger range at higher prices. In fact, they never are in a position to choose between these alternatives: they are offered either the one or the other, but never both. (Kaldor 1935, p. 50)

The presence of monopolistic competition thus provides an argument for government intervention aimed at restricting the number of firms if 'excess capacity' would be 'too large'. Paine (1936) argued that such an intervention is likely to raise prices while lowering product variety. Intervention is therefore probably disadvantageous for consumers. However, he did not take into account the income effect that would occur via higher profits.

Chamberlin, as the champion of monopolistic competition, was definitely not worried about the 'waste' associated with monopolistic competition. Rather, he points at the importance of product variety in welfare analysis.

... human beings are individuals, diverse in their tastes and desires, and moreover widely dispersed spatially. Insofar as demand has any force as a guide to production, one would expect entrepreneurs to appeal to them in diverse ways, and thus to render the output of the economy correspondingly heterogenous, using this term in the broadest sense to embrace not only the qualitative aspects of the product itself, but also the conditions surrounding its sale, including spatial location. And since what people want (...) is the starting point in welfare economics, their wants for a heterogenous product would seem to be as fundamental as everything could be. (Chamberlin 1950, p. 86)

The fundamental difference between the perspectives of Chamberlin and Kaldor is that where Kaldor only considers the individual consumer, Chamberlin considers a group of consumers. Kaldor correctly points out that the consumer normally is not given a choice between 'quantity and diversity' and that is therefore difficult to make the trade-off. If, however, the trade-off could be made and the individual consumer would choose 'quantity', this does not solve the problem of 'excess capacity'. Following the line of Chamberlin, the heterogeneity in preferences found in a group of consumers still puts a value on product diversity even when each person would only consume a single type of product. The trade-off between 'quantity and diversity' is therefore a question of the preferences of a group rather than of an individual, which makes a solution for 'optimal diversity' even more difficult, if not impossible, to obtain. The argument provided by Chamberlin does point in the direction that an excess of diversity is not likely to occur.

Chamberlin's perspective has become the dominant perspective amongst economists, especially in the fields of international trade and economic growth as will become evident in the remaining part of the section. This is not to say that interest in the trade-off between quantity and diversity has vanished (see, for example, Calvo-Armengol and Zenou 2002).

## 2.3 The Dixit-Stiglitz approach to monopolistic competition

The article "Monopolistic competition and optimum product diversity" by Dixit and Stiglitz, published in 1977, has contributed immensely to the popularity of monopolistic competition, although their results were not too different from those of Spence (1976). The framework presented by Dixit and Stiglitz (henceforth DS) is fairly simple and more general than the Hotelling spatial model. Where in Hotelling's model each consumer buys only one type of product, this is not required in the DS framework. Instead, they start by observing that a desirability for variety is implied by any objective function that has a convex indifference surface for all products. Replicating the example given in DS, "a consumer who is indifferent between the quantities (1,0) and (0,1) of two commodities prefers the mix (1/2,1/2) to either extreme." (p. 297).

The DS approach comprises both the case of a single consumer with a 'taste for variety' as well as Hotelling's result that for a given quantity of consumption, a larger number of sellers raises aggregate consumer welfare. The perspective of DS is not that of a structural model in which the preference for variety arises endogenously, but rather that of a reduced form model that takes this as an implicit assumption (Dixit 2000).

Before discussing their contribution in more detail, I will first treat the fundamental principles underlying their analysis (Dixit and Stiglitz 1993). Suppose a firm produces a single type of good, and that this firm is the only producer of this good. All firms have identical cost structures and identical market power. The price for any product,  $p$ , depends on quantity provided by the firm,  $x$ , and on the number of firms operating in the market,  $n$ .

$$p = D(x, n) \tag{2.1}$$

The profit maximization problem of the firm can be stated as  $\max_x \{D(x, n)x - C(x, n)\}$ , where  $C$  is the cost function of the firm.<sup>4</sup> The first

order condition for profit maximization is

$$D_x(x, n)x + D(x, n) - C_x(x, n) = 0. \quad (2.2)$$

After some manipulation, the price can be shown to equal marginal costs plus a markup.

$$D_x(x, n) \frac{x}{p} + 1 = \frac{C_x(x, n)}{p} \Rightarrow p = \frac{1}{\varepsilon_{px}(x, n) + 1} C_x(x, n) \quad (2.3)$$

Here,  $\varepsilon_{px}$  is the elasticity of the own price with respect to own supply. The markup over marginal costs is positive if the firm is able to influence the price of its own product by changing its supply ( $\varepsilon_{px} < 0$ ). This result is known as the Amoroso-Robinson condition (Amoroso 1930, Robinson 1950 [1933]). The common assumption made in this context is that  $\partial x_j / \partial x_i \simeq 0$  for all  $j \neq i$ , such that the firm does not take into account the reactions of the other firms and the impact thereof on the product's price and costs. Additionally, the choice of  $x_i$  is assumed not to affect demand via its influence on the budget of the buyer(s). This mode of behavior is consistent with profit maximization if the number of firms is 'large' (Dixit and Stiglitz 1993).

A second requirement – besides profit maximization – is that of free entry. When all firms are symmetric, free entry implies that the profits of all firms are zero.

$$px = C(x, n) \quad (2.4)$$

Equations 2.1, 2.3, and 2.4 should, in principle, provide sufficient information to be able to solve for  $p$ ,  $x$ , and  $n$ .

In their seminal paper from 1977 Dixit and Stiglitz applied the procedure for several cases. In the most popular case they started with a utility function with a constant elasticity of substitution (CES).<sup>5</sup>

$$u = \left( \sum_{i=1}^n x_i^\alpha \right)^{\frac{1}{\alpha}}, \quad \alpha < 1 \quad (2.5)$$

The consumer maximizes utility subject to the budget constraint  $B = \sum p_i x_i$  by choosing the optimal amount of each product type. The number of products is exogenous to the consumer but choosing  $x_i = 0$  is allowed. The first order conditions are  $p_i = \lambda^{-1} u^{1-\alpha} x_i^{\alpha-1}$ , where  $\lambda$  is the Lagrange multiplier. These conditions can be used to find the budget share of good  $i$  in terms of the quantities produced.

$$\frac{p_i x_i}{B} = \frac{\lambda^{-1} u^{1-\alpha} x_i^\alpha}{\sum_j \lambda^{-1} u^{1-\alpha} x_j^\alpha} = \frac{x_i^\alpha}{\sum_j x_j^\alpha} \quad (2.6)$$

Multiply both sides with  $B/x_i$  to get the inverse demand function as a function of the quantities and the (exogenous) budget.

$$p_i = \frac{x_i^{\alpha-1}}{\sum_j x_j^\alpha} B \quad (2.7)$$

The elasticity of the price of good  $i$  with respect to the amount supplied is  $\alpha - 1$  as the firm considers the budget of the buyer to be constant. With common constant

marginal costs given by  $c$ , the price includes a constant markup over marginal costs.

$$p = \frac{1}{\varepsilon_{px} + 1} c = \frac{c}{\alpha} \quad (2.8)$$

The assumption that all firms have a common price elasticity and marginal cost causes the price and quantity to be identical for all products. Production is assumed to require a fixed cost  $a$  such that the zero profit condition becomes

$$px = cx + a \quad (2.9)$$

The equations 2.7, 2.8, and 2.9 are the CES counterparts of equations 2.1, 2.3, and 2.4. Because all functional forms are now explicitly defined, it is possible to solve for  $x$  and  $n$  ( $p$  has already been found). After substitution for  $p$ , the equilibrium level of production is  $x = a / (\frac{1-\alpha}{\alpha} c)$ . Using the solutions for  $p$  and  $x$  in the inverse demand function yields a solution for  $n$ .

$$p = \frac{x^{\alpha-1}}{nx^{\alpha}} B \Rightarrow n = \frac{B}{px} = (1 - \alpha) \frac{B}{a} \quad (2.10)$$

The number of products is positively related to the market power of the firms and to the size of the fixed costs relative to the budget. A firm that has a large influence on the price of its product has a higher markup and does not need to sell as many products to recover the fixed costs as a firm that has a lower markup. When fixed costs can be recovered easily, many firms can enter the market before profits become negligible.

The CES utility function allows for a relatively straightforward analysis of the trade-off between quantity and diversity. In the symmetric case, the utility function can be written as  $n^{(1/\alpha)-1} (nx)$ , where the first term is the contribution of variety to utility and the second term is the contribution of quantity to utility. It is easy to see that if the production of an additional variety would not bring about additional costs ( $a = 0$ ), then utility can always be increased by the introduction of a new product. The immediate implication of this observation is that when working with CES functions, any expansion of the set of products has to be costly, otherwise utility would be infinite in 'equilibrium'.

The feature of the DS framework that the positive influence of product variety does not 'wear off' as the number of products increases, has been criticized by Pettengill (1979). According to Pettengill any increase in the number of firms will strengthen the competition between firms and therefore  $\varepsilon_{px}$  should be close to zero when  $n$  is large. This is certainly a valid criticism for Hotelling type of models as in these models a larger number of firms implies more similar products, but Pettengill's objection does not apply when new goods possess characteristics that are entirely new – an assumption implicitly made when using CES functions.

The tools provided by DS have stimulated research in many different fields, most notably in the fields of international trade and economic growth. Krugman (1979, 1980) was the first to apply the DS framework to trade theory.<sup>6</sup> Krugman's application provided an explanation for the widely observed fact that a large part of international trade is between countries with a similar endowment of production factors. Economies of scale induced by fixed costs at the location of production make it attractive for firms to concentrate their production in a single country and to export their products to the other countries in stead of setting up a plant in each country. By this mechanism, the DS framework helped to explain why similar goods are frequently traded in opposite directions.

After Krugman's papers many theorists adopted and extended his model. In fact, this body of literature has become so large that it has been labeled 'new trade theory'. Two of the most important contributions I will mention here. Ethier (1982) shifted the focus from consumption to production. In his model a single consumption good is produced from a range of intermediate goods; the production process is modeled as a CES function. Another influential publication has been the book "Market structure and foreign trade" by Helpman and Krugman (1985) in which chapters 6 to 11 cover a wide range of issues related to product differentiation and international trade.

The analysis of Chamberlin and Dixit and Stiglitz showed that without economies of scale but with free entry, there is no bound on the number of firms in the economy.<sup>7</sup> Up till 1985 the common way of preventing this unrealistic outcome had been to assume either an u-shaped average cost curve or some fixed costs of production. However, in 1985 Judd (Judd 1985b) took a different approach by restricting entry rather than by turning to economies of scale. The barrier to entry in Judd's (dynamic!) model is that each product has to be invented before it can be produced. The invention of a single new product requires  $b^{-1}$  researchers such that  $\dot{n} = bL_R$ , where  $L_R$  is the number of researchers and  $\dot{n}$  is the speed at which the number of products increases. Each new invention is protected from imitation by a patent, such that the inventor is rewarded with a (temporary) monopoly. The invention of new product types proceeds until the (expected) value of a patent has become lower than the costs of research – the wage bill of the researchers involved.

One important conclusion to be drawn from Judd's model is that with sufficient protection of patents, "product variety will grow at the same rate as population" (p. 577). This conclusion bears great similarity with the observation of DS that the number of goods is determined by the size of the fixed costs relative to the market. Both types of models predict that larger economies should have a greater variety of goods and, consequently, that their inhabitants should be richer. Chapter 8 of this dissertation contains a model that unites both approaches.

Although Judd's model contains all the ingredients for a model linking economic growth with product variety (because essentially it is one!), the major breakthrough of monopolistic competition in the theory of economic growth came five years later with a contribution by Romer (1990).<sup>8</sup> Romer's model differs from that of Judd in three respects. First, Romer considered variety in intermediates rather than consumer goods, just like Ethier (1982) did. A second difference is that he provided a more detailed specification of the microeconomic foundations of the model (this he borrowed from Grossman and Helpman 1989). The third, and the most important, point in which Romer's model deviates from Judd's is that in Romer's model previous inventions reduce the costs of current research efforts. These so-called 'knowledge spillovers' were modeled by including  $n$  on the right-hand side of the invention equation:  $\dot{n} = nbL_R$ .

This property of the model ensured that sustained economic growth is possible even without a growing population. In Judd's model the number of researcher has to increase in order to keep the growth rate of product variety ( $\dot{n}/n$ ) constant, while this is not required in Romer's model. In fact, Jones noted that in the latter model a growing number of researchers causes the growth rate of product variety to rise over time (Jones 1995a). Jones showed that this acceleration effect will diminish over time for the cases between Judd and Romer, which can be modeled as  $\dot{n} = n^\beta bL_R$  with  $0 < \beta < 1$ . The role of intertemporal knowledge spillovers in economic growth is discussed in detail in chapter 4 of this dissertation.

## 2.4 Other theories involving product variety

The Dixit-Stiglitz approach to monopolistic competition has not been limited to theories of international trade and economic growth. In an influential paper, Krugman (1991) showed how monopolistic competition can lead to agglomeration effects. It seems fair to say that this paper has revived the field of economic geography although the originality of Krugman's ideas is the subject of debate (Brakman and Garretsen 2003). Krugman and Venables (1995) and Puga and Venables (1999), amongst others, have studied how agglomeration effects of this kind might affect the convergence in per capita income between countries. A recent contribution is by Mossay (2003).

Theories aiming to explain product differentiation have not been limited to the domain of monopolistic competition. I will briefly mention some interesting cases. One example is the contribution by Stokey (1988). She presented a model of economic growth in which new product types were discovered through 'learning by doing' instead of through what can be called 'research projects' as is customary in models similar to those of Judd (1985b) and Romer (1990). This form of costless invention allowed Stokey to preserve perfect competition: instant imitation by competitors does not prevent research because research is not required to be profitable. Inventions are simply assumed to be inseparable from ordinary productive activities. One interesting feature of her model is that the invention of new products gradually leads to the extinction of old ones. More recently, Greenwood and Uysal (2004) have used a model with perfect competition to estimate the welfare gains from rising product variety in the United States between 1800 and 2000.

Another approach can be found in the literature on vertical product differentiation (Gabszewicz and Thisse 1980, Sutton 1986). Vertical product differentiation means that products are only differentiated with respect to quality, and not with respect to function (this would be horizontal differentiation). Although derived from Hotelling's model, the mechanisms of models with vertical differentiation differ substantially from those of models with monopolistic competition. First of all, products with a higher quality are assumed to be more costly to produce. This feature is important because when products are identically priced, only the product with the highest quality will be sold. The number of products in these models very much depends on the structures of preferences and production costs. Flam and Helpman (1987) have applied vertical product differentiation to a model of North-South trade.

The last example that I will mention is the literature on multiproduct firms, which has its roots in the literature on vertical product differentiation (Shaked and Sutton 1982, Shaked and Sutton 1990). Here, product variety is regarded as an instrument of competing oligopolists. By producing some optimal number of types, a monopolist may or may not be able to prevent the introduction of new products by an outside firm.

## Notes

1. While Robinson defined perfect competition as "... a state of affairs in which the demand for output of an individual seller is perfectly elastic." (Robinson 1934, p. 104), Chamberlin was more inclined to equate perfect competition just with free entry.
2. The two authors were quite eager to engage in semantic battles. Robinson (1934) tried to explain Chamberlin what perfect competition is, while Chamberlin (1937) tried to educate Robinson on the difference between imperfect and monopolistic competition.
3. The implicit assumption typically made is that the average cost curve is u-shaped such that there is an unique and finite level of production that minimizes average costs.
4. In stead of choosing an optimal level of production, a firm may also set an optimal price.
5. Contrary to the claim by DS (p. 298), Spence (1976, pp. 226-233) had also used the CES utility function in his analysis.
6. A similar approach was developed by Dixit and Norman (1980), probably simultaneously with Krugman.
7. For an early discussion of entry and economies of scale see Triffin, 1949, pp. 117-123 and 145-157.
8. In an earlier paper, Romer (1987) presented a model of economic growth based on DS with a quadratic cost function for intermediates and with free entry.





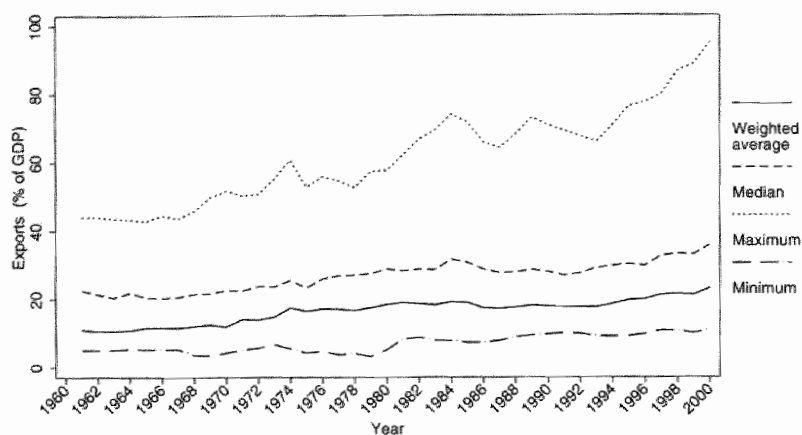
## Empirical trends

This chapter presents the empirical developments that provide the context of the rest of the thesis. These empirical trends are divided in three categories. Section 3.1 treats the rapid growth of the value of trade, the convergence of per capita income in the OECD, and the relation between per capita income and population size. This section introduces the context for the model of economic integration presented in chapter 8. The (recent) developments regarding education and research are the subject of section 3.2. The models of chapter 7 provide a theoretical framework that links these trends together. Section 3.3 introduces the data on product variety that are used throughout this dissertation. A more detailed empirical study of product variety can be found in chapter 10. At the end of the chapter, a brief overview is given of the most important empirical trends.

### 3.1 Trade, income convergence, and scale

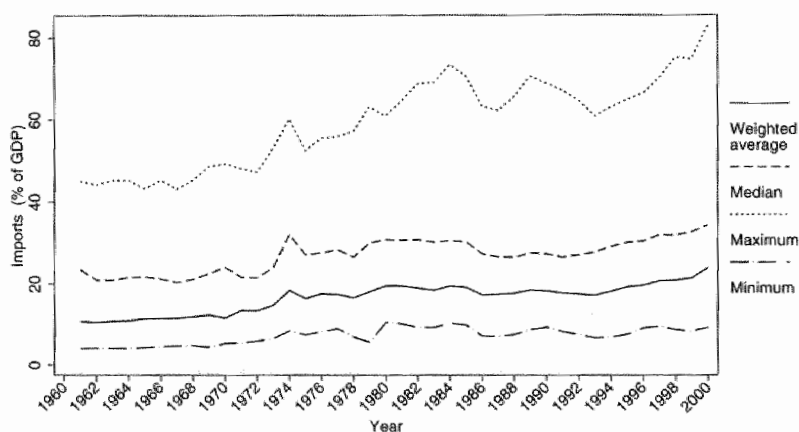
The empirical analysis on economic growth customarily starts with the assumption that every country has a more or less independently functioning economy.<sup>1</sup> For high-income OECD countries, however, this assumption is not even approximately correct. Figure 3.1 shows how the share of exports in GDP for a group of 24 OECD countries evolves over time. The data for figure 3.1 have been taken from the World Development Indicators database (WDI; compiled by the World Bank) and from the Penn World Tables 6.1 (PWT) (Heston, Summers, and Aten 2002). The solid line is the total value of exported goods and services by all 24 countries divided by their aggregated GDP. This line indicates that for every dollar of OECD production in 2000, twenty cents crossed a (any) border. This was ten cents in 1960. If international income transfers would have been included, these figures would have been even higher. The dashed line in figure 3.1 is the median export share for the same group of countries. This line lies above the aggregate export share, indicating that small countries rely more on exports than large countries. The minimum and maximum export share are also plotted.

The value of imports as a percentage of GDP is plotted in figure 3.2, which is comparable to figure 3.1. The similarity of the two figures is not surprising as foreign currency earned with exports is usually not hoarded but used for the purchase of foreign goods and services. Conversely, spending more than is being earned with exports leads to an accumulation of debts, which cannot be considered



**Figure 3.1.** Exports of goods and services expressed in percentage of GDP (OECD)

Sources: World Bank (WDI); Penn World Table 6.1

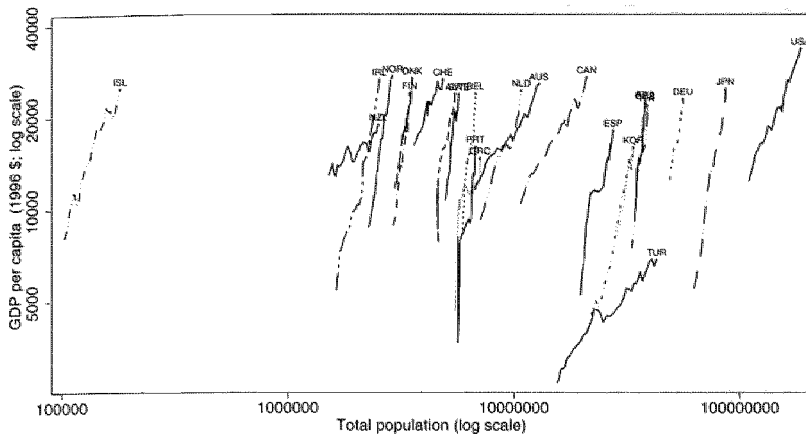


**Figure 3.2.** Imports of goods and services expressed in percentage of GDP (OECD)

Sources: World Bank (WDI); Penn World Table 6.1

a sustainable strategy.<sup>2</sup> Figures 3.1 and 3.2 convincingly show that, at least for this group of countries, trade matters – and a lot!

The assumption of independent national economies (INE) is extremely important for the evaluation of theories of economic growth. To give an example, the models of Dixit and Stiglitz (1977), Judd (1985b), and Jones (1995a) discussed in the previous section all predict that population size determines the standard of living in an economy. According to INE, these theories can be tested empirically by verifying that large countries are richer than small countries. Figure 3.3 is a scatterplot that relates GDP per capita to population size for the period 1970 to 2000 (a list explaining the codification of the countries is provided in table A.1).<sup>3</sup>



**Figure 3.3.** GDP per capita and population between 15 and 64 years of age (OECD; 1960-2000)

Sources: Penn World Table 6.1; World Bank (WDI)

Although the United States is both the largest and the richest country, there is no obvious relation between the size of a country and its production per head. The relation between real GDP per capita and population size seems to be of an intertemporal rather than cross-sectional nature. This convergence of production per capita in the OECD has been linked to the rise of international trade by, amongst others, Ben-David (1993), Williamson (1996), and Ventura (1997).

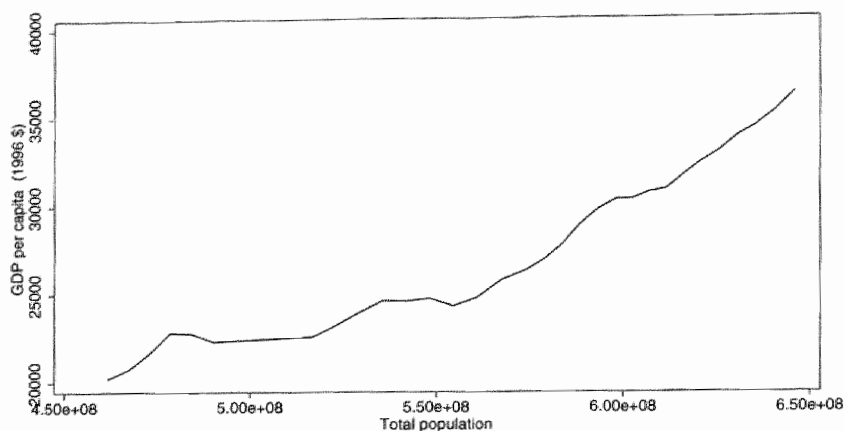
The lack of a relation between country size and GDP per capita suggests that either the models by Dixit and Stiglitz, Judd, and Jones are widely of the mark, or that INE does not hold. Given the importance of trade (figures 3.1 and 3.2) it seems sensible to drop INE (further evidence is provided in chapter 10).

Figure 3.4 again displays GDP per capita on the vertical axis and population size on the horizontal axis, but now the group of 24 OECD members is taken as the unit of analysis. Looking at figure 3.4, the models of Dixit and Stiglitz, Judd, and Jones suddenly seem to be much more plausible than figure 3.3 suggests. However, figure 3.4 does not prove the existence of a direct relation between the size of the population and GDP per capita: the correlation might be spurious. A study by Kremer (1993) quite convincingly showed that in the (very) long run, standards of living depend on the size of the population (and vice versa).

When testing theories of economic growth, the assumption that countries have economies that function independently can be very influential. An example of this has been given above and more examples will follow. In this dissertation, empirical analysis will start from the assumption that the members of the OECD together constitute a single economy. This sample of countries has been chosen for two reasons. First, these 24 countries form the core of the world economy. Second, for some variables no reliable data is available for non-OECD countries.

## 3.2 Education and research

Given the prominent role of technology for the standard of living, education is likely to be important as well. It can be seen from figure 3.5 that the average

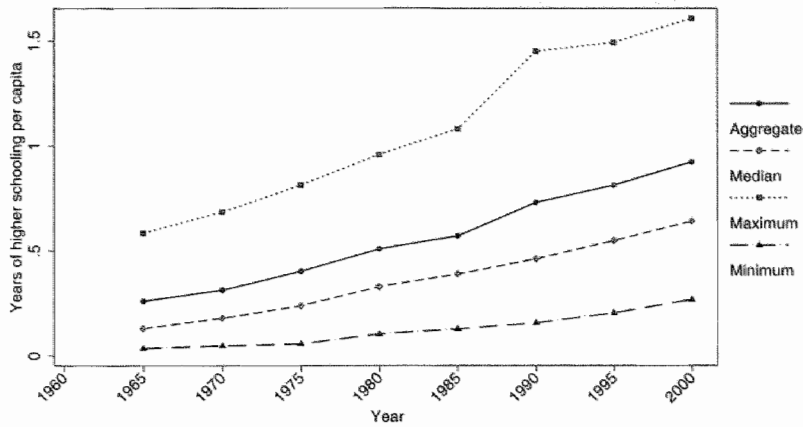


**Figure 3.4.** GDP per capita and population between 15 and 64 years of age in the OECD (1970-2000)

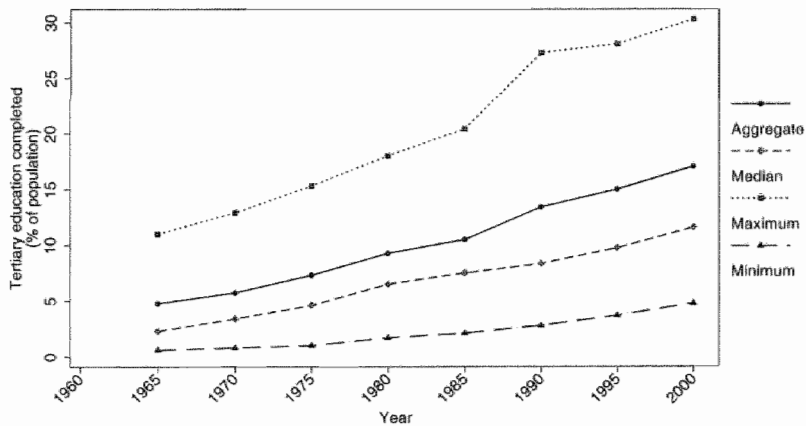
Sources: Penn World Table 6.1; World Bank (WDI)

number of years of higher schooling has risen steadily in the OECD. The same can be said of the share of the population that has completed tertiary education, which is plotted in figure 3.6.<sup>4</sup> At the same time, figure 3.7 shows that the distribution of schooling has become less unequal and more uniform.<sup>5</sup> The data used in figures 3.5, 3.6, and 3.7 stem from the dataset compiled by Barro and Lee (2000, 2001).

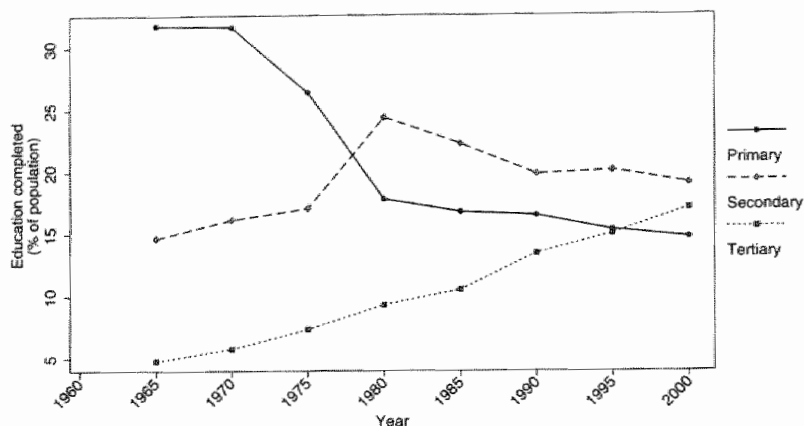
Whether the rise in educational attainment coincides with an increase in research expenditure, is not entirely clear. In figure 3.8, the share of gross expenditure on research and development (GERD) in GDP is plotted against time. The data are taken from the OECD's Main Science and Technology Indicators database (MSTI). Because of their irregular availability, the data have been inter- and, occasionally, extrapolated. Although the fraction of aggregate R&D expenditure in GDP seems to remain constant over time, the maximum, median, and minimum values exhibit an upward trend. This suggests that the poor availability of data may have biased the series on aggregate expenditure downwardly. A second hint in this direction is provided by figure 3.9, which shows the proportion of researchers in the workforce. In this figure both the median and aggregate series are increasing. The data on the proportion of researchers in the workforce also stem from the OECD's MSTI database.



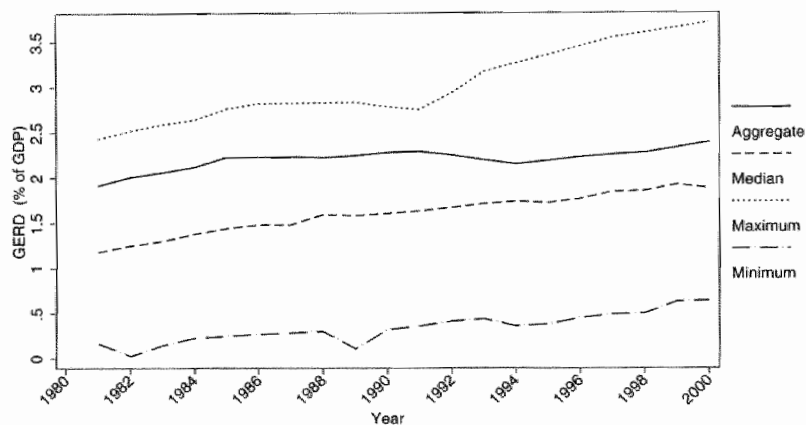
**Figure 3.5.** Years of higher schooling per capita (OECD)  
Sources: Barro and Lee (2000); World Bank (WDI)



**Figure 3.6.** Proportion of population above 25 that has completed tertiary education (OECD)  
Sources: Barro and Lee (2000); World Bank (WDI)

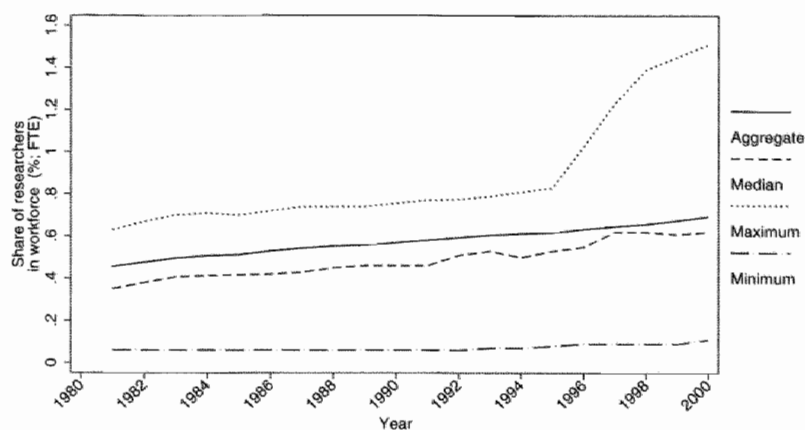


**Figure 3.7.** Proportion of population above 25 with primary, secondary, or tertiary education as the highest level of education completed (OECD)  
Sources: Barro and Lee (2000); World Bank (WDI)

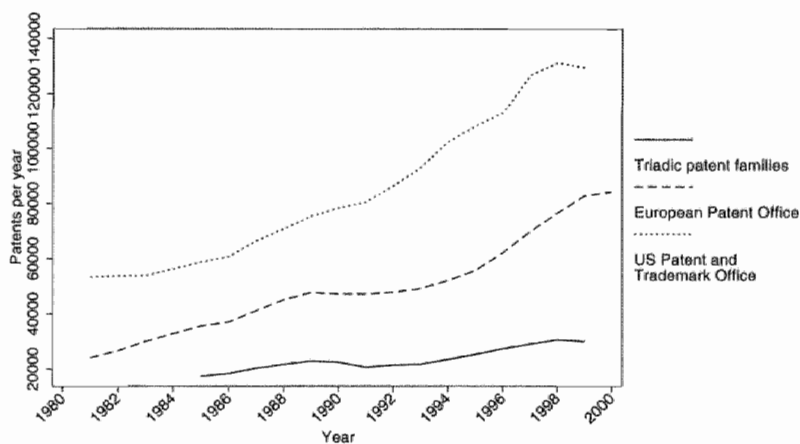


**Figure 3.8.** Gross expenditure on R&D (GERD) expressed in percentage of GDP (OECD); data have been interpolated at the country level  
Sources: OECD (MSTI); Penn World Tables 6.1

Data on newly registered/granted patents give another indication that research is becoming a more prominent economic activity over time. Figure 3.10 displays the number of new patents granted by the United States Patent and Trademark Office (USPTO), the number of new patents registered with the European Patent Office (EPO), and the number of 'triadic patent families'. A patent family is a group of patents on a single invention issued by multiple countries. A patent family is 'triadic' if patents have been issued by USPTO, EPO, and the Japanese Patent Office (JPO).<sup>6</sup> It is clear from figure 3.10 that all three patent flows have risen between 1981 and 1999. Interestingly, the growth rates of the three series are not identical, especially the number of new triadic patent families is growing relatively slowly.

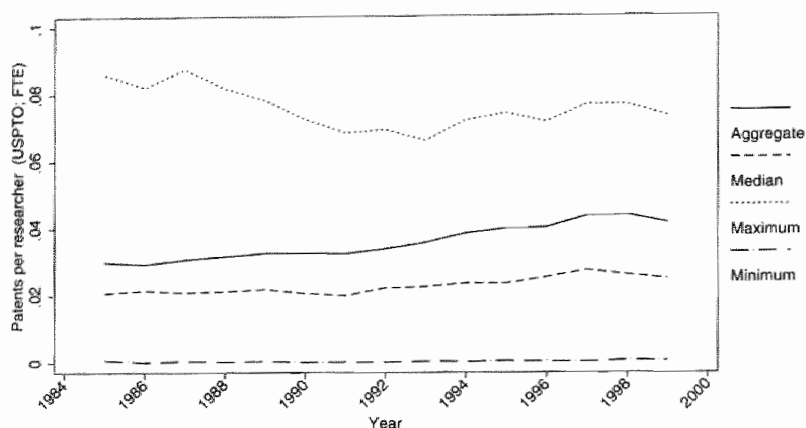


**Figure 3.9.** Number of researchers (FTE) in percentage of total employment; data have been interpolated at the country level  
Sources: OECD (MSTI); World Bank (WDI)



**Figure 3.10.** Patent flows: triadic families, European patents, and US patents (OECD)  
Source: OECD (MSTI)





**Figure 3.11.** Number of USPTO patents per researcher (FTE; OECD); data have been interpolated at the country level  
Source: OECD (MSTI)

Figures 3.5 to 3.10 demonstrate that the population of the OECD has become more highly educated and that the importance of research as an economic activity is rising over time. It is not unlikely that these two trends are related: an abundant supply of highly educated workers may very well be a stimulant for research. There might also be other mechanisms at work, such as the positive influence of new technology on the productivity of researchers.

The productivity of researchers is at the centre stage of the literature on endogenous growth. (section 2.3 and chapter 4). One empirical fact that has received a lot of attention is the surge in the number of researchers in the United States. Jones, for example, pointed out that the rapid growth in the number of researchers in the United States, France, Germany, and Japan coincided with a constant rate of total factor productivity growth in these countries (Jones 1995b, pp. 517-518). This implies that the marginal productivity of researchers is declining. Kortum (1997, p. 1390) emphasized the fall in the number of patents per researcher in the United States.

Dropping the INE assumption yields a different perspective. As can be seen from figure 3.11, the number of USPTO patents per researchers in the OECD has risen only moderately between 1985 and 1999. How can the difference between this finding and those of Kortum and Jones be explained? First of all, the rise of research in the United States could have come at the expense of research in the (smaller) countries of the OECD.<sup>7</sup> Second, the definitions of “researcher” are different. The data used by Jones and Kortum originate from the National Science Foundation (NSF), which simply counts the number of researchers. The data used in figure 3.11 come from the OECD and are in full-time equivalents (FTE). A shift from full-time to part-time research positions thus can explain part of the difference. Furthermore, the NSF data only pertain to the private sector, whereas the OECD data also include the public sector.

### 3.3 Product variety

Technological change is quite intimately linked to product variety. An early example of this relationship is provided by table 3.1, which contains the share in total production of the industrial sector in 18th century Britain. This period has been labeled the (British) Industrial Revolution because of the rapid pace of invention and adoption of new technologies. The data are from Crafts and Harley (1992) who used them for the computation of their 'revised best guess index of industrial production'. Although the data are probably not extremely accurate, it can safely be concluded that the number of large sectors has risen substantially: from 9 sectors around 1700 to 22 a century later.

The rise in the number of newly registered patents illustrated by figure 3.10 suggests that contemporary product variety is increasing as well. Direct evidence on the number of produced types is not available because of a lack of detailed production data.<sup>8</sup> Fortunately, this lack of production data can be compensated by trade data. Detailed and internationally comparable data on imports and exports available from 1961 onwards. Figure 3.12 shows the number of imported and the number of exported product groups for which a positive value is recorded in the database.<sup>9</sup> A separate plot is displayed for each of the 24 countries. The data stem from the OECD's International Trade by Commodity Statistics (ITCS) and UNSD's COMTRADE databases (details are available in appendix A). Products are classified according to revision 2 of the SITC system for the period 1961 to 1990 and according to revision 3 from 1991 onwards.

A striking characteristic of the graphs in figure 3.12 is the stepwise increases in both series. According to the OECD these 'jumps' have to do with the way in which the products are classified. It is not only the actual classification of products in the database that matters, but it is also the classification of the original data from which the database is constructed that is influential. The changes in the (underlying) classification are clearly visible in figure 3.13, which displays the number of traded product groups for the OECD (both intra- and extra-OECD trade) together with the total value of trade recorded in the database.

The first clear break in the data occurs in 1976. This break coincides with a switch from SITC revision 1 to SITC revision 2. A small decrease in the number of groups can be observed in 1989. This decrease is a consequence of adoption of a new version of the Harmonized System (HS 88) for the classification of products in the underlying data. According to the OECD, the adoption of HS 88 made countries report a lower number of groups than before because of confidentiality issues. The last break occurs in 1991 when the products in the database started to be classified according to SITC revision 3.

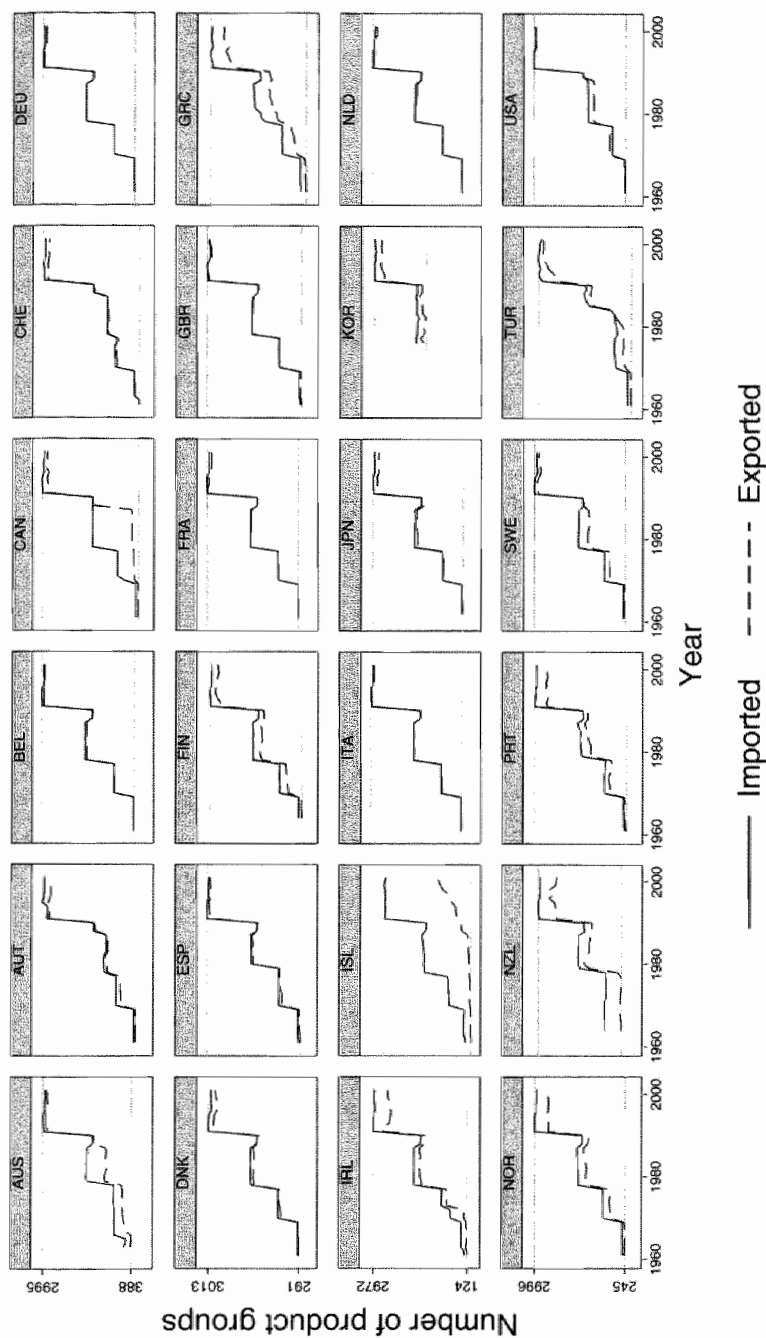
The breaks in the number of traded product types did not lead to discontinuities in the value of trade. This suggests that each revision of the classification system is merely a disaggregated version of the previous revision, such that entirely new product categories are not captured. However, whether the changes in the number of products reflect administrative improvements or actual changes in product variety, cannot be concluded from the data. If administrative improvements are the primary cause of the breaks, then these breaks have to be filtered out. Figure 3.14 displays the evolution of the number of groups if SITC revision 3 would have been in use since 1961, assuming all discontinuities are merely a consequence of administrative improvements. After filtering out classification changes, the number of groups is constant over time.<sup>10</sup>

Table 3.1. Relative sectoral size during the British Industrial Revolution

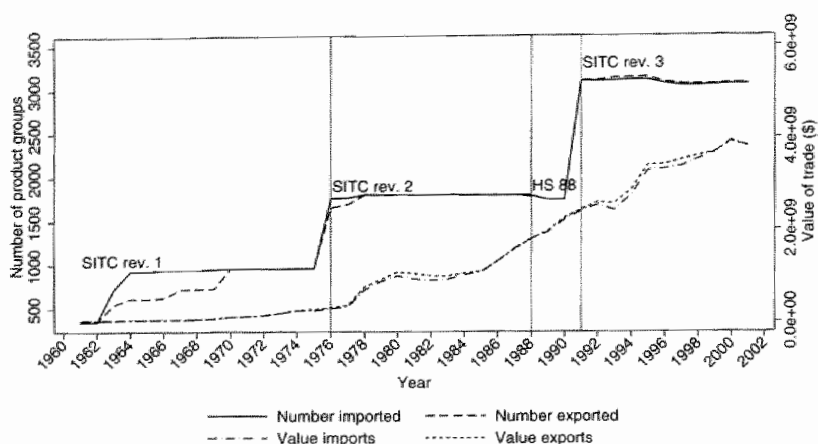
Begin period	1700	1711	1713	1720	1722	1727	1739	1761	1771	1780	1787	1789
End period	-1710	-1712	-1719	-1721	-1726	-1738	-1760	-1770	-1779	-1786	-1788	-1800
Beer	18.4	17.1	16.3	13.8	8.3	8.3	5.8	5.8	5.8	5.8	5.8	5.7
Candles		6.9	6.6	5.6	3.4	3.4	2.3	2.3	2.3	2.3	2.3	2.2
Coal	15.1	14.1	13.4	11.3	6.8	6.7	4.6	4.6	4.5	4.5	4.5	4.4
Copper goods									0.5	0.5	0.5	0.5
Copper ore							0.5	0.5	0.4	0.4	0.4	0.4
Cotton	11.6	10.8	10.3	8.7	5.3	5.2	2.8	6.7	6.7	6.7	6.7	6.7
Hides and skins					39.6	39.5	27	23.2	23.1	23.1	23.1	22.5
Iron				15.6	9.5	9.4	6.5	6.5	6.5	6.5	6.5	6.5
Linen	27.9	26.0	24.8	20.9	12.6	12.5	8.6	8.6	8.5	8.5	8.5	8.3
Malt	5.8	5.4	5.1	4.3	2.6	2.6	1.8	1.8	1.8	1.8	1.8	1.7
Paper			1.2	1.0	0.6	0.6	0.4	0.4	0.4	0.4	0.4	0.4
Ships												
Silk	15.1	14.1	13.4	11.3	6.8	6.8	4.6	4.6	4.6	4.6	5.0	4.4
Silk cloth												
Silk yarn											3.5	3.0
Soap			3.7	3.1	1.9	1.9	1.3	1.3	1.3	1.3	1.3	1.4
Sugar	1.0	0.9	0.8	0.7	0.4	0.4	0.3	0.3	0.3	0.3	0.3	0.3
Tin	2.3	2.1	1.9	1.6	0.9	0.9	0.6	0.6	0.6	0.6	0.6	0.6
Tobacco	2.8	2.6	2.5	2.1	1.3	1.2	0.9	0.9	0.9	0.9	0.9	0.9
Wool							32	31.9	31.8	31.8	31.8	31.0
Wool cloth										17.4	17.4	17.0
Wool yarn										14.4	14.4	14.0

Sectoral value added in % of total value added

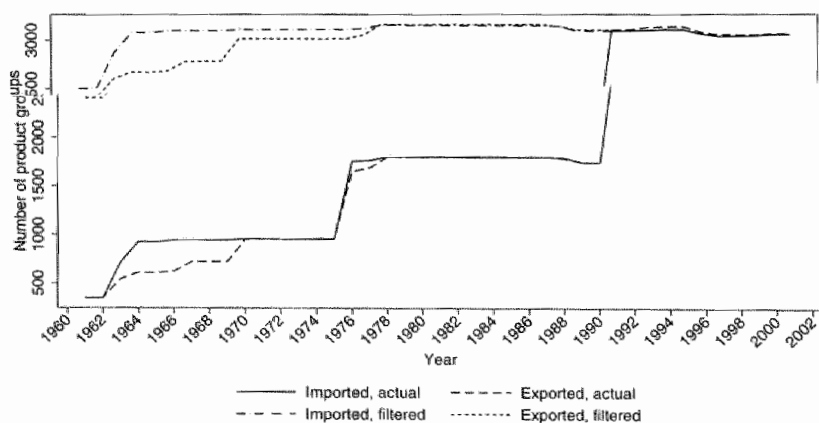
Source: Crafts and Harley (1992, table A3.2)



**Figure 3.12.** Number of traded product groups by country (1961-2001)  
Sources: OECD (ITCS), UNSD (COMTRADE)



**Figure 3.13.** Number of traded product groups in the OECD (1): changes in classification and the recorded value of trade  
Sources: OECD (ITCS), UNSD (COMTRADE)



**Figure 3.14.** Number of traded product groups in the OECD (2): filtering out changes in classification  
Sources: OECD (ITCS), UNSD (COMTRADE)

### 3.4 Summary

Table 3.2 displays the growth rates of a selection of the variables that have been discussed above. The growth rates are obtained by regressing the logarithm of the variables reported in the first on time and a constant. The standard deviations of the growth rates are shown in the third column. The periods to which the estimates apply and the data sources can be found in the fourth and fifth column, respectively.

The general picture that arises from table 3.2 is that the growth of GDP per capita has coincided with an increase in the importance of trade, education, and research, with population growth, and with a rise in the number of product types, although this latter conclusion is open to debate because of data issues. The theoretical consequences of the changes in the level and distribution of educational attainment are discussed in section 7.2. The consequences of the growth in international trade are treated in chapter 8. The growth of product variety will be analyzed in more detail in chapter 10.

**Table 3.2.** Trends in selected macroeconomic variables for the OECD

Variable	Yearly growth	Standard deviation	Period	Source
Population	.012	.0005	1960-2001	WDI
Workforce	.014	.0005	1960-2001	WDI
GDP PPP per capita	.037	.001	1960-2000	PWT
Export (% of GDP)	.017	.001	1960-2000	WDI, PWT
Import (% of GDP)	.017	.001	1960-2000	WDI, PWT
Years of higher schooling per capita	.037	.002	1960-2000	BL, WDI
Tert. education compl. (% of pop.)	.037	.002	1960-2000	BL, WDI
R&D expenditure*†	.006	.001	1981-2000	MSTI, PWT
Median R&D exp.*†	.023	.001	1981-2000	MSTI
Researchers* (FTE; % of employment)	.020	.0005	1981-2000	MSTI, WDI
Patents per researcher* (USPTO; FTE)	.023	.002	1981-1999	MSTI
Imported variety	.039	.002	1964-2001	app. A
Imported variety (filtered)	-.0002	.0002	1964-2001	app. A
Exported variety	.047	.003	1964-2001	app. A
Exported variety (filtered)	.003	.0006	1964-2001	app. A
Flow of triadic patents	.036	.004	1985-2000	MSTI
Flow of EPO patents	.059	.003	1981-2000	MSTI
Flow of USPTO patents	.056	.002	1981-1999	MSTI

The reported growth rate is a deterministic trend fitted to the logarithm of the variable

\* Data have been interpolated at the country level

† percentage of GDP

## Notes

1. Hypotheses on economic growth are typically tested using regression analysis over a cross-section of countries, considering each country to be a separate economy. For a survey of the literature see Barro and Sala-i-Martin (2004, chap. 12).
2. The United States did not seem to be bothered much by a high current account deficit during the 1990s and the beginning of the second millennium.
3. No data on GDP PPP per capita is available for Germany prior to 1970.
4. It is striking to see that the aggregate OECD values for education lie above the median values. This indicates that larger countries tend to have a better educated population than small countries.
5. The reader might have noticed that the sum of the shares is about 60%. This is due to the fact that the data refer to the level of education that has been completed rather than attained. For 'education attained' the shares add up to approximately 100%. The trends for 'education attained' are similar to those for 'education completed'.
6. The use of triadic patent families has two advantages – in theory at least. First, unimportant inventions are not likely to be registered in all three countries, such that triadic patent families can be considered to relate to significant inventions. Second, the institutional and geographical bias is less when patents are issued by several countries. In practise, however, not many non-Japanese firms have their inventions patented in Japan because of the high costs. Patenting an invention with all three offices is also not really necessary when production is subject to economies of scale. A patent valid for a part of the market is usually sufficient to prevent imitation.
7. Jones noted that as the growth in the number researchers has also been large in France, Germany, and Japan, choosing the OECD as the unit of analysis should not lead to different conclusions (Jones 1995b, p. 519).
8. Detailed production data are available for a selection of European countries and for the states of the US (see appendix A). These data are only available for a small number of years, and have been used in the cross-sectional analysis of chapter 10.
9. A 'product group' is defined as a collection of, presumably similar, product types.
10. For individual countries the growth rate in the number of product groups is positive (see table 10.4 and section 10.4 on the curse of classification).





# Knowledge spillovers and population growth: Endogenous growth theory 1995-2003

Chapter 2 ended with a brief discussion of the models by Judd (1985a), Romer (1990), and Jones (1995a). This discussion briefly touched upon the importance of intertemporal knowledge spillovers for the relation between population size and product variety. The current chapter provides a technical survey of the recent theoretical literature on economic growth, which centers on the joint effects of intertemporal knowledge spillovers and population growth on research activities and economic growth. The survey will rely heavily on a generalized model of economic growth that will be laid out in the subsequent sections. Several of the results that are derived in this chapter will be referred to in subsequent chapters in order to avoid repetition of algebra.

## 4.1 Introduction

Two kinds of R&D-driven models of economic growth have been devised in the first half of the 1990s. In the first group of models, economic growth was the result of an expansion in the number of products. The most important examples of this ‘love-of-variety’ approach are Romer (1990) and Grossman and Helpman (1991a, chap. 3). The second group relied on improvements in the quality of existing product types as a source of economic growth. The contributions by Grossman and Helpman (1991c, 1991b, 1991a, chap. 4) and Aghion and Howitt (1992) belong to this second group. This approach to modelling economic growth is commonly known as the ‘quality ladder’ or ‘creative destruction’ approach.

In a pair of influential papers, Jones (1995a, 1995b) criticized the R&D-driven growth models of the early 1990s on grounds of their empirical plausibility. A common feature of the early growth models is that they yield a constant rate of economic growth when the size of the population is constant. Jones noted that when the population is not constant but instead is growing over time, these models predict explosive economic growth, meaning that the growth rate increases without bound. He argued that this is an extremely unrealistic feature because the rapid rise in the number of researchers in the United States did not coincide with an acceleration of total factor productivity growth.<sup>1</sup>

Jones (1999) illustrated his criticism with the following example. Suppose that output,  $Y$ , depends on the amount of labor involved in production,  $L_Y$ , and on the stock of useful ideas,  $A$ .

$$Y = A^\alpha L_Y \quad (4.1)$$

The stock of ideas symbolizes either the number of different intermediate products or their average quality. New ideas are discovered using  $L_A$  researchers and the stock of previously generated ideas.

$$\dot{A} \equiv \frac{dA}{dt} = A^\beta L_A \quad (4.2)$$

A dotted symbol denotes a flow. The early growth models had  $\beta = 1$ , which implies that a constant number of researchers is sufficient to ensure a constant growth rate of ideas.

$$\hat{A} \equiv \frac{\dot{A}}{A} = L_A \quad (4.3)$$

A variable wearing a hat is a growth rate. This specification, however, is troublesome as a growing number of researchers causes the growth rate of the stock of ideas to rise over time. If  $g_L$  is the growth rate of the population and  $L_A/L_Y$  is constant, then taking growth rates at both sides of the equation yields  $\hat{\hat{A}} = g_L$ . This acceleration in the accumulation of ideas has far-reaching consequences for production. The growth rate of output per capita is given by  $\hat{Y} - g_L = \alpha \hat{A}$ . If  $\hat{A}$  is growing over time, then so is the growth rate of output per capita.

Jones argued that the requirement  $\beta = 1$  yields knowledge spillovers that are too strong. In particular, he showed that  $0 \leq \beta < 1$  is sufficient to rule out explosive economic growth. In order to see this, first divide both sides of the R&D-equation by  $A$  and take growth rates, assuming  $\hat{L}_A = g_L$ .

$$\hat{\hat{A}} = (\beta - 1) \hat{A} + g_L \quad (4.4)$$

As  $\hat{\hat{A}}$  is a decreasing function of  $\hat{A}$ , the growth rate of the stock of ideas eventually becomes constant. Such an equilibrium is commonly referred to as the 'steady state' of the model. Note that this equilibrium depends on the assumption  $\hat{L}_A = g_L$ , which is equivalent with a constant ratio  $L_A/L_Y$ . The steady state growth rate of  $A$  can be found by having  $\hat{\hat{A}} = 0$  and solving for  $g_A$ . (Steady state growth rates are denoted by  $g_X$ , where  $X$  is the variable of interest.)

$$g_A = \frac{1}{1 - \beta} g_L \quad (4.5)$$

Steady state per capita output growth is now given by

$$g_Y - g_L = \alpha g_A = \frac{\alpha}{1 - \beta} g_L. \quad (4.6)$$

The conclusion to be drawn from this result is that, as long as  $\beta < 1$ , the growth rate of per capita output is constant, even when the population is growing. As in this case economic growth ultimately depends on population growth, it is considered to be 'semi-endogenous' rather than 'endogenous'. Jones argued that requiring  $\beta < 1$  does not only yield more plausible predictions than  $\beta = 1$ , but that the former specification is also to be preferred on theoretical grounds. In particular,

$\beta < 1$  can be considered to be a more general case than the 'knife-edge' condition  $\beta = 1$ . Kortum (1997) and Segerstrom (1998), taking the quality ladder approach, have presented more elaborate versions of Jones' model.

The response to Jones' semi-endogenous growth model arrived in 1998 with the contributions by Young (1998), Peretto (1998a), Dinopoulos and Thompson (1998), and, somewhat later, Howitt (1999). These authors were uncomfortable with the outcome of Jones' model that population growth is vital for economic growth and consequently they came up with models that are characterized by a constant steady state rate of economic growth – whether the population is growing or not. I will label this kind of economic growth 'neo-endogenous'. A feature common to all neo-endogenous growth models is that they allow for simultaneous variety and quality research. The nature of knowledge spillovers, however, is different for both types of research. Only weak spillovers are assumed for research aimed at the invention of new product types, such that sustained variety expansion requires population growth. At the same time, strong spillovers are assumed for research on quality improvements, enabling sustained quality growth in the absence of population growth. These assumptions are crucial for the reconciliation of endogenous growth with population growth.

In a nutshell, the mechanism that avoids explosive rates of economic growth works as follows. A growing variety of products raises the number of quality researchers that are required for the growth rate of average quality to remain constant. For example, if the number of products increases from 10 to 20, the number of quality researchers has to be doubled in order to maintain the same number of quality researchers per product type. The weak spillovers in variety research ensure that this 'dilution' effect only occurs when the population is growing. The aggregate number of quality researchers is constant when the size of the population is fixed, while it is growing whenever the population is growing.

Li (2000) responded to the neo-endogenous growth models with the critique that these models rely heavily on restrictive assumptions about the nature of spillovers between variety and quality research. Li argued that for models with both variety and quality research, semi-endogenous growth is still more general than (neo-)endogenous growth. Like endogenous growth, neo-endogenous growth depends on knife-edge assumptions.

Peretto and Smulders (2002) reacted to the neo-endogenous growth models in exactly the opposite way. Contrary to Li they proposed a model in which spillovers are stronger instead of weaker than customary in neo-endogenous growth models. The feature that enables neo-endogenous growth despite strong spillovers is the assumption that spillovers are confined to a network of researchers working on related issues. The number of networks is related to the scale of the economy, such that a larger economy does not automatically imply that spillovers are larger for individual researchers. In particular, Peretto and Smulders assumed that a constant fraction of new inventions brings along a new field of research. A consequence of this specification is that for large economies the number of networks grows at the same rate as the number of products.

The remainder of this chapter provides a detailed treatment of the developments sketched above. This will be done by means of a generalized model of economic growth. The advantage of this approach is that it allows for straightforward comparison of the various models. The disadvantage is that generalization requires a model to have many features, which makes it difficult to analyze. In order to reduce complexity, I have chosen to leave out some of the features present in some of the models discussed above. For example, I do not

consider capital to be an input for research like Howitt (1999) and Segerstrom (2000) do, nor do I consider the stochastic nature of inventions.<sup>2</sup>

The structure of the model is explained in sections 4.2 to 4.5. Section 4.6 treats the conditions under which the model generates endogenous growth in the spirit of the early growth models. Semi-endogenous growth is the topic of section 4.7 and neo-endogenous growth is discussed in section 4.8. The chapter ends with some concluding remarks on the relation between the size of the population and the number of products.

## 4.2 Production

Two different frameworks are used for modelling the production of ordinary goods. In the first framework a variety of intermediate products are being ‘assembled’ by workers into a homogenous consumption good. In the second framework there is a variety of consumer goods that do not require intermediate inputs. The model introduced below allows for the analysis of both frameworks, but it will be discussed as a model with a variety of intermediate products.

Production proceeds in two stages. First, intermediates are produced with the use of labor only. Each of these product types is produced by a separate firm, such that the intermediate sector is characterized by monopolistic competition. The produced amount for intermediate type  $i \in N$  is

$$x(i) = bl(i), \quad (4.7)$$

with  $b$  being a constant and  $l(i)$  the amount of labor employed by intermediate firm  $i$ . For analytical convenience, the set of intermediate types  $N$  is assumed to be continuous. The size of the set  $N$  is indicated by the measure  $n$  (the mass of types), which is equivalent to the number of product types if the set  $N$  would have been discrete rather than continuous.

In the second stage of production, the intermediate goods are assembled into homogenous final products,  $y$ , with the use of labor,  $l_y$ . There are  $n_y$  symmetric firms in the final goods sector and competition is perfect. The production process of a firm  $j$  that produces final goods is described by

$$y(j) = l_y(j)^{1-\delta} \left[ \int_N q(i)^\nu x(ij)^\gamma di \right]^{\frac{\delta}{\gamma}}, \quad (4.8)$$

where  $0 < \gamma < 1$ ,  $0 < \delta \leq 1$ , and  $x(ij)$  is the quantity supplied by  $i$  to  $j$ . The intermediate-specific parameter  $q(i)$  indicates the quality (or ‘usefulness’) of intermediate good  $i$ . The current setup is a generalization of a class of models of economic growth in the sense that these models can be recovered by imposing restrictions on the parameters  $\delta$  and  $\nu$ . For example, if  $\delta = 1$ ,  $y$  can be interpreted as a consumption index over a variety of consumer goods. Table 4.1 gives an overview of the configurations that have been used in the more recent literature.

The profits of firms in the final goods sector can be expressed as

$$\pi_y(j) = l_y(j)^{1-\delta} \left[ \int_N q(i)^\nu x(ij)^\gamma di \right]^{\frac{\delta}{\gamma}} - \int_N p(i) x(ij) di - wl_y(j), \quad (4.9)$$

where final goods are taken to be the numeraire,  $p(i)$  is the price of intermediate  $i$ , and  $w$  is the wage rate. Profits are maximized by choosing the optimal amounts

**Table 4.1.** Parameterization of production

$\delta = 1$	$\nu = \nu$	Peretto (1998a, 1998b, 2003b) Peretto and Smulders (2002)
	$\nu = \gamma$	Young (1998) Jones (1999)
	$\nu = 1 - \gamma$	Dinopoulos and Thompson (1998) Li (2000) Thompson (2001)
$\delta = \gamma$	$\nu = 1$	Howitt (1999) Segerstrom (2000)

of  $l_y$  and  $x(ij)$ .<sup>3</sup> The corresponding first order conditions (foc's) are

$$\frac{\partial \pi_y(j)}{\partial x(ij)} = \delta y(j)^{1-\frac{\gamma}{\delta}} q(i)^\nu x(ij)^{\gamma-1} - p(i) = 0 \quad (4.10)$$

$$\frac{\partial \pi_y(j)}{\partial l_y(j)} = (1 - \delta) \frac{y(j)}{l_y(j)} - w = 0. \quad (4.11)$$

The firms in the final goods sector can have any size compatible with perfect competition. In order to overcome this indeterminacy, it assumed that all  $n_y$  firms are of the same size. Rewriting these foc's gives the inverse demand functions for intermediate goods and the wage bill in terms of output, respectively.

$$p(i) = \delta y^{1-\frac{\gamma}{\delta}} q(i)^\nu x(i)^{\gamma-1} n_y^{1-\gamma} \quad (4.12)$$

$$wl_y = (1 - \delta) y \quad (4.13)$$

Intermediate firms also maximize their profits. The profits of firm  $i$  are given by  $\pi(i) = p(i, x)x(i) - wl(i)$  subject to  $x(i) - bl(i) = 0$ . Because each product is produced by only one firm, the firm is able to influence the price by the amount that is being produced, that is why the firm considers the price to be a function of the quantity produced. The associated Lagrangian function is

$$\mathcal{L}(i) = \max_{x, l} \{p(i, x)x(i) - wl(i) + \lambda_x(i)(x(i) - bl(i))\} \quad (4.14)$$

with foc's

$$\frac{\partial \mathcal{L}(i)}{\partial x(i)} = \frac{\partial p(i)}{\partial x(i)} x(i) + p(i) + \lambda_x(i) = 0 \quad (4.15)$$

$$\frac{\partial \mathcal{L}(i)}{\partial l(i)} = -w - \lambda_x(i)b = 0 \quad (4.16)$$

From these foc's and from the inverse-demand function (equation 4.12) follows that the profit-maximizing price is a mark-up over marginal costs (see also section 2.3).

$$p(i) = \frac{w}{\gamma b} = p \quad (4.17)$$

As marginal costs are identical across firms, prices are identical as well. Notice that the price of a good does not depend on its quality, even though competition

is monopolistic. Insert the solution for the price into the inverse demand function to obtain the partial equilibrium solution for the quantities of intermediate production.

$$x(i) = (\delta\gamma b)^{\frac{1}{1-\gamma}} w^{\frac{1}{\gamma-1}} y^{\frac{\gamma-\delta}{\delta(\gamma-1)}} n_y q(i)^{\frac{\nu}{1-\gamma}}. \quad (4.18)$$

The appearance of  $n_y$  in this last expression reveals a weak spot of the 'variety of intermediate goods'-approach to modelling economic growth. Unless  $\delta = 1$ , the number of producers in the final sector matters for the production of intermediate firms. This fact has been overlooked in the literature, which is characterized by the implicit modelling assumption  $n_y = 1$ . This omission is important. If  $n_y$  would be allowed to vary, there is no assurance that producers in the final goods sector respond to price changes in the intermediate goods sector. Changes in intermediate prices could then be fully absorbed by entry or exit of final good producers. Hence, the assumption made in equation 4.10 that  $x(ij)$  depends on  $p(i)$  is incorrect if  $n_y$  is not constant. The spirit of the result that intermediate producers are able to charge a price higher than marginal costs, however, is not invalidated as they still have a monopoly over their own product types. What remains unspecified are the properties of the demand curve faced by an intermediate firm. In this chapter, I will follow the literature by assuming  $n_y = 1$ .<sup>4</sup> In the other chapters, I will limit the analysis to the 'variety in consumption goods'-approach, which assumes  $\delta = 1$ .

The solution for the price can also be used to simplify the expression for the profits of intermediate firms.

$$\begin{aligned} \pi(i) &= px(i) - wl(i) = \left(p - \frac{w}{b}\right) x(i) \\ &= \left(\frac{1}{\gamma} - 1\right) \frac{w}{b} x(i) = \left(\frac{1}{\gamma} - 1\right) wl(i) \end{aligned} \quad (4.19)$$

After substitution of  $x(i)$  and  $n_y$  this becomes

$$\pi(i) = \left(\frac{1}{\gamma} - 1\right) b^{\frac{2-\gamma}{1-\gamma}} (\delta\gamma)^{\frac{1}{1-\gamma}} w^{\frac{\gamma}{\gamma-1}} y^{\frac{\gamma-\delta}{\delta(\gamma-1)}} q(i)^{\frac{\nu}{1-\gamma}} \quad (4.20)$$

A firm can only change its profits by changing the quality of its products. The next section discusses how firms can do this.

### 4.3 Research

The probability that a researcher discovers a quality improvement is  $\phi_q dt$ . Such a quality improvement raises  $q(i)$  with a fixed amount  $\theta$ . Quality research is assumed to be performed 'in-house', meaning that the firm producing the product does the research itself. This has more empirical support than the 'creative destruction'-kind of quality improvements (Malerba and Orsenigo 1995, Malerba, Orsenigo, and Peretto 1997) and allows for 'cumulative patents' (patents that complement each other). At every point in time, each firm sets up a quality research project that lasts an infinitesimal period. Research projects are funded by issuing shares in the research project (or by issuing options on the patents resulting from research).

The number of researchers employed in such a research project are denoted by  $l_q(i)$ . The number of researchers that makes a discovery during  $dt$  is

$$\int_0^{l_q(i)} \phi_q dt dl = l_q(i) \phi_q dt \text{ w.p.1.} \quad (4.21)$$

(w.p.1. stands for 'with probability one'.) The total increase in the quality of good  $i$  per unit of time equals the number of inventions per unit of time, multiplied by the size of a single improvement,  $\theta$ .<sup>5</sup>

$$\dot{q}(i) = l_q(i) \phi_q \theta \text{ w.p.1} \quad (4.22)$$

As before, a dot above a symbol denotes its derivative with respect to time. The research project sells the flow of inventions at price  $v_q(i)$  and distributes these revenues among all the researchers that have participated in the project, regardless of their contribution.

Similarly, the probability that a researcher discovers a new product is  $\phi_n dt$ . The flow of new products is

$$\dot{n} = l_n \phi_n \text{ w.p.1,} \quad (4.23)$$

where  $l_n$  is the number of researchers that independently seek to invent entirely new products. Each new product has an initial quality that is randomly drawn from the distribution of the quality levels of existing products. Research is funded through the sales of new patents.

It is possible that some researchers independently discover the same type of improvement. If this occurs, one discovery is (randomly) labeled as successful while the other discoveries are labeled as being unsuccessful and are ignored. The arrival rate of successful inventions for a single researcher takes into account the probability of duplicate inventions and could therefore be negatively related to the amount of researchers that participate in the research project. This is reflected by the following specification of the arrival rate.

$$\phi_j = \mu_j l_j^{-\varphi_j}, \quad j \in \{q, n\} \quad (4.24)$$

The parameter  $\mu \geq 0$  scales the relation between  $\phi$  and the other variables;  $\varphi \geq 0$  determines the size of the duplication effect.

The number of researchers working on a project is not the only factor affecting the arrival rates of inventions – previous research results also matter. Knowledge, the product of researchers, differs from ordinary products in two important respects (Romer 1990). First, knowledge is non-rival, meaning that its availability or scarcity is not affected by the intensity with which it is used. One implication of non-rivalry is that once a new product is invented it can be produced on any scale without additional R&D costs. The non-rivalry of knowledge thus leads to economies of scale as average R&D costs per unit of production decline with increases in the scale of production. Second, knowledge is non-excludable meaning that it is difficult, if not impossible, to use knowledge without compromising its secrecy. The non-excludability of knowledge implies that imitation is less costly than invention. Furthermore, once a new product is brought to the market it might inspire other people to produce similar products.

These two factors, non-rivalry and non-excludability, allow researchers to benefit from the results of others. In models of economic growth this phenomenon is incorporated by letting the arrival rate of inventions, or, equivalently, the productivity of researchers, depend on the stock of previous inventions.<sup>6</sup> The 'spillovers' from the stock of inventions to the productivity of researchers has been modelled in roughly four ways.



- **Leading edge:** the highest quality level affects the arrival rate of quality improvements for all products.
- **Average quality level:** the average quality level across products affects the arrival rate of a quality improvement.
- **Number of products:** the current number of product affects the arrival rates of new products and quality improvements.
- **Knowledge base:** some measure of the size of a knowledge base affects the arrival rates of new products and quality improvements.

**Leading edge** The leading edge approach is used in conjunction with stochastic quality changes, something that has been neutralized in the framework of this chapter. Stochastic quality changes allow for competition within the same industry in the form of ‘patent races’: the firm with the best product is able to make a profit until another firm invents a better product. In models of this kind the firm that produces the highest level of quality faces a weaker incentive to innovate than the lagging firms because a leading firm does not have a more advanced firm to imitate and because any new product type will make the current type obsolete. A typical arrival rate for quality improvements is

$$\phi_q = \mu_q l_q^{-\varphi_q} \max_{i \in N} \{q_i\}^{v_q}, \quad (4.25)$$

where  $v_q$  is a parameter indicating the influence of the leading edge. Early contributions following this approach are Grossman and Helpman (1991c) and Aghion and Howitt (1992). More recently, Segerstrom (2000) has also presented a model with this kind of spillovers. Young (1998) uses a similar functional specification but then with  $v_q < 0$  in order to simulate that new inventions are more complex and therefore more difficult to discover.

**Average quality level** The models that rely on spillovers from the leading firm are not very transparent. The alternative assumption that knowledge spillovers depend on the average quality level yields simpler models and is therefore more popular. With this approach the arrival rate can be formulated as

$$\phi_q = \mu_q l_q^{-\varphi_q} \left( \frac{\int_N q_i di}{n} \right)^{\varepsilon_q}, \quad (4.26)$$

where  $\varepsilon_q$  determines the influence of the average level of quality. This approach has been adopted by Peretto (1998a, 1998b, 2003b), Segerstrom (2000), and Thompson (2001).

**Number of products** Besides the average or maximum level of quality, also the variety of products has been modelled as a source of knowledge spillovers, both for research aimed at the invention of entirely new types of products and for research aimed at quality improvement. The arrival rates belonging to this approach are

$$\phi_j = \mu_j l_j^{-\varphi_j} n^{-\chi_j}, \quad j \in \{q, n\} \quad (4.27)$$

This specification has been used by Li (2000).

**Knowledge base** A fourth perspective on modelling knowledge spillovers is to let arrival rates depend on a ‘knowledge base’, a set of ideas. Suppose that there are  $n(k)$  firms doing research in network  $k$ . The number of different networks is  $K$ . All firms belonging to the same network are supposed to have access to each other’s knowledge. The body of knowledge available in a network is the result of previous research done by the firms in that network. If the set of ideas needed to produce good  $i$  with quality level  $q(i)$  is denoted  $Q(i)$ , then the total body of knowledge of network  $k$  is given by

$$Q(k) = \bigcup_{N(k)} Q(i), \quad (4.28)$$

where  $N(k)$  is the set of firms operating in  $k$ . Suppose further that the quality level is related to its knowledge base by  $|Q(i)| = q(i)^{\eta_q}$  with  $\eta_q \geq 0$ .<sup>7</sup> If all knowledge is product specific, that is if  $Q(i) \cap Q(j) = \emptyset$  for all  $i \neq j$ , then the measure of the total knowledge base is  $|Q(k)| = \int_{N(k)} |Q(i)| di = \int_{N(k)} q(i)^{\eta_q} di$ . However, it is more likely that a single idea is used with the production of several goods. As a result the measure of total knowledge is less than the sum of the measures of the knowledge bases of the individual products, or  $|Q(k)| < \int_{N(k)} |Q(i)| di$ . The proportion of ideas that are used on several occasions rises with the size of the knowledge base. Incorporating knowledge duplication, the measure of total knowledge in a network is given by the CES-like aggregate of quality levels

$$|Q(k)| = \left[ \int_{N(k)} q(i)^{\eta_q} di \right]^{\eta_Q} \quad (4.29)$$

with  $0 \leq \eta_Q \leq 1$ .

The arrival rates of inventions are assumed to depend positively on the measure of total knowledge in the network.

$$\phi_j = \mu_j l_j^{-\varphi_j} |Q(k)|^{\varepsilon_j}, \quad 0 \leq \varepsilon_j \leq 1, \quad j \in \{q, n\} \quad (4.30)$$

All ideas currently available in the network are equally likely to provide inspiration for a quality improvement or for the invention of a new product. The network approach to knowledge spillovers in models of economic growth is due to Peretto and Smulders (2002). They also added the knowledge base of the firm itself to the equation in order to be able to vary the degree of excludability of knowledge.

$$\phi_j(i) = \mu_j l_j^{-\varphi_j} |Q(k)|^{\varepsilon_j} |Q(i)|^{\zeta_j}, \quad j \in \{q, n\} \quad (4.31)$$

The four approaches outlined above can be united by synthesizing equations 4.25 to 4.31 into a single – admittedly slightly awkward – equation.

$$\phi_j(i) = \mu_j l_j^{-\varphi_j} \frac{|Q(k)|^{\varepsilon_j}}{n(k)^{\chi_j}} |Q(i)|^{\zeta_j} \max_{i' \in N(k)} \{|Q(i')|\}^{\nu_j} \quad (4.32)$$

$$= \mu_j l_j^{-\varphi_j} \frac{\left[ \int_{N(k)} q(i)^{\eta_q} di \right]^{\eta_Q \varepsilon_j}}{n(k)^{\chi_j}} q(i)^{\eta_q \eta_Q \zeta_j} q_{\max(k)}^{\eta_q \eta_Q \nu_j} \quad (4.33)$$

Table 4.2 displays the various configurations of parameters that have been used in the recent literature. Not all models fit into this classification seamlessly. For example, some models rely on the stochastic nature of invention for their results (Young 1998), while others have capital goods as inputs to R&D (Howitt 1999,

Table 4.2. Parameterization of research and knowledge spillovers

	$\eta_q$	$\eta_Q$	$\varphi_q$	$\varepsilon_q$	$\chi_q$	$\zeta_q$	$v_q$	$\varphi_n$	$\varepsilon_n$	$\chi_n$	$\zeta_n$	$v_n$	$K^*$
Young (1998)	0	0	0	0	0	1	-1	0	0	0	1	-1	1
Dinopoulos and Thompson (1998)	1	1	0	1	1	0	0	0	0	0	0	0	1
Peretto (1998a, 1998b)	1	1	0	1	1	0	0	0	0	0	0	0	1
Jones (1999)	1	1	0	1	0	0	0	(0,1)	0	0	0	0	1
Howitt (1999)	1	1	0	0	1	0	0	(0,1)	(0,1)	0	0	-1	1
Segerstrom (2000)	1	1	(0,1)	1	1	0	(0,1)	(0,1)	1	1	0	<0	1
Li (2000)	$\frac{\gamma}{1-\gamma}$	>0	0	>0	<1	<0	0	0	$\varepsilon_n$	$\chi_n$	0	0	1
Thompson (2001)	1	1	(0,1)	1	1	0	0	0	0	0	0	0	1
Peretto and Smulders (2002)	1	1	0	(0,1)	0	$1-\varepsilon_q$	0	0	>0	0	$-\varepsilon_n$	0	$K^*$
Peretto (2003b)	1	1	(0,1)	1	1	0	0	0	0	0	0	0	1

\*  $K = K_0 + \kappa(n - n_0)$

Segerstrom 2000). Nevertheless, the table gives an idea of the principal structure of the models under consideration.

As quality improvements are effectively non-random for individual products, and as new products have quality levels that are randomly drawn from the distribution of the quality levels of existing products, all firms can be treated as being symmetrical. Hence, arrival rates become

$$\phi_j = \mu_j l_j^{-\varphi_j} n^{\varepsilon_{nj}} q^{\varepsilon_{qj}} K^{\varepsilon_{kj}} \quad (4.34)$$

$$\varepsilon_{nj} \equiv \eta_Q \varepsilon_j - \chi_j, \quad \varepsilon_{qj} \equiv \eta_Q \eta_Q (\varepsilon_j + \zeta_j + v_j), \quad \varepsilon_{kj} \equiv \chi_j - \eta_Q \varepsilon_j,$$

with  $n(k) = n/K$ . Both quality-levels and employment are identical for all firms and consequently the quality level of any product grows at the rate

$$\hat{q} = l_q^{1-\varphi_q} \theta \mu_q n^{\varepsilon_{nq}} q^{\varepsilon_{qq}-1} K^{\varepsilon_{kq}}. \quad (4.35)$$

The number of products grows at the rate

$$\hat{n} = l_n^{1-\varphi_n} \mu_n n^{\varepsilon_{nn}-1} q^{\varepsilon_{qn}} K^{\varepsilon_{kn}}. \quad (4.36)$$

These last two equations are of central importance for studying the effects of knowledge spillovers on economic growth. Before we proceed with the analysis of these effects, a digression is made on how households determine the amount and type of research activities they are willing to fund.

## 4.4 Households

Households are assumed to optimize the net present value of instantaneous utility,  $u$ . Instantaneous utility is derived purely from the household's consumption of final goods,  $c$ . It is not necessarily optimal to consume all income as it possible to use a part of the income to buy shares in research activities. This can be an attractive option because the owners of shares receive dividends. Households are facing the problem of planning their consumption path such that the discounted sum of instantaneous utility is maximal, while taking into account the size of the household,  $l(j)$ .

$$\max_{\{c\}_{t=0}^{\infty}} \left\{ \int_0^{\infty} l(j) u(c_t(j)) \exp[-\rho t] dt \right\} \quad (4.37)$$

Instantaneous utility is discounted at rate  $\rho$ , otherwise any non-decreasing stream of instantaneous utility would yield infinity in the expression above. The intuitive interpretation of this discounting is, of course, that households care more about the present than about the future.

Total income has to equal total expenditure plus savings. Household income consists of wages, interest on loans, dividends stemming from research on new products, and dividends stemming from quality research.

$$wl(j) + rz(j) + \int_N \frac{m(ij)}{M} \pi(i) di. \quad (4.38)$$

In this equation  $r$  is the interest rate,  $z$  is the amount lent by household  $j$  to other households,  $m(ij)$  is the number of shares in firm  $i$  possessed by household  $j$ , and  $M$  is the number of shares per firm.

Household expenditure comprises consumption, new loans to other households (or to firms), and the purchase of shares. In order to prevent speculation, and thus

share-price bubbles, each share can be traded only once. There is no second-hand market for shares. Only shares can be bought that are currently issued. A part of the shares is issued by research projects aiming to invent new products, another part is issued by existing firms for the funding of a quality improving research project. Household expenditure is the sum of all these amounts.

$$l(j)c(j) + \dot{z}(j) + \frac{m_n(j)}{M}\dot{n}v_n + \int_N \frac{m_q(ij)}{M}v_q(i)di \quad (4.39)$$

$v_n$  is the value of a patent for a new product,  $v_q$  is the value of a quality research project. The budget constraint of the household equates income with expenditure.

$$wl(j) + rz(j) + \int_N \frac{m(ij)}{M}\pi(q(i))di = \quad (4.40)$$

$$l(j)c(j) + \dot{z}(j) + \frac{m_n(j)}{M}\dot{n}v_n + \int_N \frac{m_q(ij)}{M}v_q(i)di$$

It is important to note that the allocation of rents stemming from quality patents is rather arbitrary when research is done 'in house'. A new quality improving patent brings about a rise in profits but these extra profits are not automatically appropriated by the shareholders of the new patent. This problem can be avoided by assuming symmetric (identical) households. I will call a single household from the set of symmetric households a representative household. The representative household is indifferent to the allocation of profit streams to the patents that are used to produce a good as this does not alter the total amount of dividends it receives. Effectively, the assumption of a representative household implies the addition of the restriction<sup>8</sup>

$$\frac{m_n(ij)}{M} = \frac{m_q(ij)}{M} = \frac{l}{L} = h^{-1} \quad (4.41)$$

to the optimization problem, where  $h$  is the number of households. The growth rate of  $l$  is assumed to equal that of the total population,  $g_L$ . This implies that the number of households is fixed.

The representative household cannot be allowed to lend to other households as this would destroy symmetry. Rewrite the budget constraint (4.40) for the representative household.

$$wl + h^{-1} \int_N \pi(q(i))di = lc + h^{-1}\dot{n}v_n + h^{-1} \int_N v_q(i)di \quad (4.42)$$

Because firms are symmetric and new firms start with the quality level of existing firms, this expression can be simplified even further.

$$wl + h^{-1}\Pi(n, q) = lc + h^{-1}\dot{n}v_n + h^{-1}nv_q \quad (4.43)$$

$$\Pi(n, q) \equiv n\pi(q)$$

Differentiation of  $\Pi$  with respect to time yields

$$\dot{\Pi} = \pi\dot{n} + n \left( \frac{\partial \pi}{\partial q}\dot{q} + \frac{\partial \pi}{\partial t} \right), \quad (4.44)$$

where the partial derivative of the individual profit rate with respect to time comprises all changes in dividends that cannot be influenced by a single household.

All funds acquired by means of issuing shares are used to pay the wages of researchers. For quality research this follows because otherwise there would simply

be a redistribution of wealth among shareholders, which would be neutral in the case of symmetric households.<sup>9</sup> For research aimed at discovering new products, this follows from the free entry and exit of researchers. The 'zero-profit' conditions for research are

$$v_q = w l_q \quad (4.45)$$

$$\dot{v}_n = w l_n. \quad (4.46)$$

Using the first 'zero-profit' condition we find that  $\dot{q} = v_q \phi_q / w$ .

Taking the above into consideration, the intertemporal optimization problem becomes

$$\max_{\{c\}_0^\infty} \left\{ \int_0^\infty \frac{L_0}{h} u(c_t(j)) \exp[g_L - \rho t] dt \right\} \quad (4.47)$$

subject to

$$c = w + L^{-1} \Pi(n, q) - L^{-1} v_n \xi - L^{-1} n v_q \quad (4.48)$$

$$\dot{\Pi} = \pi \xi + n \left( \pi_q \frac{v_q \phi_q}{w} + \frac{\partial \pi}{\partial t} \right) \quad (4.49)$$

where  $\xi$  equals  $\dot{n}$  and  $\pi_q$  is the derivative of  $\pi$  with respect to  $q$ .

The present-value Hamiltonian for this problem is<sup>10</sup>

$$\mathcal{H} = \max_{\xi, v_q} \left\{ \frac{L_0}{h} \exp[g_L - \rho t] u \left( w + L^{-1} \Pi(n, q) - L^{-1} v_n \xi - L^{-1} n v_q \right) + \lambda \left( \pi \xi + n \left( \pi_q \frac{v_q \phi_q}{w} + \frac{\partial \pi}{\partial t} \right) \right) \right\} \quad (4.50)$$

The representative household (or rather, the representative shareholder) decides on how many patents for new products to fund by means of  $\xi$ , and on how much to spend on quality improving R&D per firm by means of  $v_q$ .

The corresponding optimality conditions are  $\partial \mathcal{H} / \partial \xi = \partial \mathcal{H} / \partial v_q = 0$  and  $\partial \mathcal{H} / \partial \Pi = -\dot{\lambda}$ . After substitution, we have

$$\frac{\partial \mathcal{H}}{\partial \xi} = -\frac{L_0}{h} \exp[g_L - \rho t] u_c L^{-1} v_n + \lambda \pi = 0 \quad (4.51)$$

$$\frac{\partial \mathcal{H}}{\partial v_q} = -\frac{L_0}{h} \exp[g_L - \rho t] u_c L^{-1} n + \lambda n \frac{\pi_q \phi_q}{w} = 0 \quad (4.52)$$

$$\frac{\partial \mathcal{H}}{\partial \Pi} = \frac{L_0}{h} \exp[(g_L - \rho) t] u_c L^{-1} = -\dot{\lambda}, \quad (4.53)$$

where  $u_c$  is the derivative of  $u$  with respect to  $c$ . The first two equations each provide a solution for the multiplier  $\lambda$ .

$$\lambda = \frac{L_0}{h} \exp[(g_L - \rho) t] u_c L^{-1} \frac{v_n}{\pi} \quad (4.54)$$

$$\lambda = \frac{L_0}{h} \exp[(g_L - \rho) t] u_c \frac{L^{-1} n}{n \frac{\pi_q \phi_q}{w}} \quad (4.55)$$

Together these two equations lead to the 'arbitrage equation' that equates the return on the different kinds of patents.

$$\frac{\pi}{v_n} = \frac{\pi_q \phi_q}{w} = \frac{\pi_q \dot{q}}{v_q} \quad (4.56)$$

Differentiate the first solution for  $\lambda$  and equate the result with  $-\partial H/\partial \Pi$  in order to get a condition for the optimal path of consumption.

$$\frac{\partial \mathcal{H}}{\partial \Pi} + \dot{\lambda} = 0 \quad (4.57)$$

$$= \frac{L_0}{h} \exp[(g_L - \rho)t] u_c L^{-1} + \lambda (g_L - \rho + \hat{u}_c - g_L + \hat{v}_n - \hat{\pi}) \quad (4.58)$$

$$-\hat{u}_c = \frac{\pi}{v_n} - \rho + \hat{v}_n - \hat{\pi} \quad (4.59)$$

From now on, the instantaneous utility function is defined here as  $u(c) = \ln c$ .<sup>11</sup> Consequently, the optimality conditions transform into

$$\dot{c} = \frac{\pi}{v_n} - \tilde{\rho} = \frac{\pi_q \dot{q}}{v_q} - \tilde{\rho}. \quad (4.60)$$

The definition  $\tilde{\rho} \equiv \rho + \hat{\pi} - \hat{v}_n$  is introduced to simplify notation later on. The last two equations are examples of what is called the 'Ramsey rule' after Ramsey (1928). Originally, this rule was derived for households that were saving in order to finance the accumulation of capital goods rather than research (Cass 1965, Koopmans 1965).

## 4.5 Aggregation

The microeconomic foundations of the model have been laid out in sections 4.2 to 4.4. The current section outlines the macroeconomic relations arising from these foundations. We will start with the macroeconomics implied by the production side of the model. Noting that  $n_y = 1$ , aggregate output of the final goods sector,  $Y$ , can easily be derived from equation 4.8.

$$Y = y = L_y^{1-\delta} n^{\frac{\delta}{\gamma}} q^{\frac{\delta\nu}{\gamma}} x^{\delta} \quad (4.61)$$

In this last equation,  $q$  and  $x$  are an intermediate firm's quality level and total production, respectively.  $L_y$  is aggregate labor in the final goods sector. The aggregate wage bill of the final goods sector and total expenditure on intermediate products follow directly from equation 4.13.

$$wL_y = (1 - \delta)Y \quad (4.62)$$

$$np_x = \delta Y \quad (4.63)$$

Symmetry in the intermediate goods sector yields  $x = bL_x/n$ , where  $L_x$  is aggregate intermediate labor. Using the last two equations in combination with equation 4.19, the revenue of the intermediate sector can be subdivided into profits and wages, respectively.

$$\Pi = (1 - \gamma)\delta Y \quad (4.64)$$

$$wL_x = \gamma\delta Y \quad (4.65)$$

Equations 4.62 and 4.65 imply that the ratio  $L_y/L_x$  equals  $\frac{1-\delta}{\gamma\delta}$ . Taking advantage of this fixed ratio, aggregate final output can be written in terms of  $L_x$ .

$$Y = \left(\frac{1-\delta}{\delta\gamma}\right)^{1-\delta} b^{\delta} n^{\frac{\delta}{\gamma}-\delta} q^{\frac{\delta\nu}{\gamma}} L_x \quad (4.66)$$

Only one aggregate relation has to be derived from the savings behavior of households. Aggregate consumption is given by  $Y = Lc$ , and, consequently, the growth rate of aggregate consumption equals the growth rate of per capita consumption plus the rate of population growth.

$$\hat{Y} = g_L + \frac{\pi}{v_n} - \bar{\rho} = g_L + \frac{\pi_q \hat{q}}{v_q} - \bar{\rho} \quad (4.67)$$

The profit-value ratio for new goods can be derived using the non-profit condition for research (equation 4.46) in combination with equation 4.20.

$$\frac{\pi}{v_n} = \frac{\left(\frac{1}{\gamma} - 1\right) w L_x}{n \frac{w l_n}{\hat{n}}} = \left(\frac{1}{\gamma} - 1\right) \frac{L_x}{l_n} \hat{n} \quad (4.68)$$

Similarly, the rate of return for quality improvements can be derived using equations 4.45 and 4.20.

$$\frac{\pi_q \hat{q}}{v_q} = \frac{\nu}{\gamma} b^{\frac{2-\gamma}{1-\gamma}} (\delta \gamma)^{\frac{1}{1-\gamma}} w^{\frac{1}{\gamma-1}} Y^{\frac{\gamma-\delta}{\delta(\gamma-1)}} q^{\frac{\nu}{1-\gamma}} l_q^{-1} \hat{q} \quad (4.69)$$

In this expression, the wage rate can be replaced with  $\gamma \delta Y / L_x$ , and  $Y$  can be substituted for using 4.66.

$$\frac{\pi_q \hat{q}}{v_q} = \frac{\nu}{\gamma} b^{\frac{2-\gamma}{1-\gamma}} L_x^{\frac{1}{1-\gamma}} Y^{\frac{\gamma}{\delta(\gamma-1)}} q^{\frac{\nu}{1-\gamma}} l_q^{-1} \hat{q} \quad (4.70)$$

$$= \frac{\nu}{\gamma} b^2 \left( \frac{1-\delta}{\gamma \delta} \right)^{\frac{\gamma(1-\delta)}{\delta(\gamma-1)}} L_x^{\frac{\gamma-\delta}{\delta(\gamma-1)}} (n l_q)^{-1} \hat{q} \quad (4.71)$$

After substitution of the rates of return into the Ramsey rule, two equations for economic growth are obtained that will prove to be useful later on.

$$\hat{Y} = g_L + \left(\frac{1}{\gamma} - 1\right) \frac{L_x}{l_n} \hat{n} - \bar{\rho} \quad (4.72)$$

$$= g_L + \frac{\nu}{\gamma} b^2 \left( \frac{1-\delta}{\gamma \delta} \right)^{\frac{\gamma(1-\delta)}{\delta(\gamma-1)}} L_x^{\frac{\gamma-\delta}{\delta(\gamma-1)}} (n l_q)^{-1} \hat{q} - \bar{\rho} \quad (4.73)$$

## 4.6 Endogenous growth

Growth models that are characterized by spillovers strong enough to ensure sustained economic growth in the absence of population growth, are called 'endogenous growth models'. For economic growth to be positive and constant, the growth rates of quality or variety have to be positive and constant as well. This follows directly from equation 4.66.

$$\hat{Y} = \left( \frac{\delta}{\gamma} - \delta \right) \hat{n} + \frac{\nu \delta}{\gamma} \hat{q} \quad (4.74)$$

When the size of the population is fixed, sustained growth in variety and quality needs to be generated by a constant number of researchers. A growing number of researchers is simply not sustainable when the total population does not grow. This rather strong condition has profound consequences for the parametrization of the spillover process. The parameter settings compatible with endogenous growth



driven by both quality improvements and variety expansion are  $\varepsilon_{nn} = \varepsilon_{qq} = \varepsilon_{nq} = 1$  and  $\varepsilon_{qn} = 0$  ( $K = 1$ ). With these settings, equations 4.35 and 4.36 can be simplified considerably:

$$\hat{q} = l_q \theta \mu_q n = \theta \mu_q L_q \quad (4.75)$$

$$\hat{n} = l_n \mu_n \quad (4.76)$$

$L_q$  is the total number of researchers working on quality improvements. The two equations above indicate that the growth rates of quality and variety are constant if the number of researchers remains constant for each type of research activity.

Another feature of these equations is that the spillovers are asymmetric. The arrival rate of quality improvements depends on the variety of products, whereas the arrival rate of the number of products does not depend on the average quality level. This asymmetry is required for a simple reason. Increasing product variety dilutes the number of quality-researchers per product and this has to be compensated for by a higher productivity of quality researchers. Setting  $\varepsilon_{nq} = 1$  causes the dilution of quality-researchers to be offset exactly by the gain in productivity that arises from spillovers. This dilution effect plays an important role with the neo-endogenous growth models discussed in section 4.8.

In order to solve for the growth rate of the economy, the allocation of labor has to be found. Equations 4.72, 4.73, 4.74, 4.75, and 4.76, provide sufficient information to solve for  $l_n$  and  $L_q$ . Let us start by substituting for  $\hat{n}$  in 4.72 using 4.76.

$$\hat{Y} = \left(\frac{1}{\gamma} - 1\right) \mu_n L_x - \tilde{\rho} \quad (4.77)$$

The Ramsey rule for the funding of quality improvements combined with 4.71 yields a similar expression.

$$\hat{Y} = \frac{\nu b^2 \theta \mu_q}{\gamma} \left(\frac{1-\delta}{\gamma \delta}\right)^{\frac{\gamma(1-\delta)}{\delta(\gamma-1)}} L_x^{\frac{\gamma-\delta}{\delta(\gamma-1)}} - \tilde{\rho} \quad (4.78)$$

A solution for  $L_x$  can simply be found by equating the two expression for  $\hat{Y}$ .

$$L_x = \left(\frac{\nu b^2 \theta \mu_q}{(1-\gamma) \mu_n}\right)^{\frac{\delta(\gamma-1)}{\gamma(\delta-1)}} \frac{\gamma \delta}{1-\delta} \quad (4.79)$$

With  $L_x$ , and thus  $L_y (= \frac{1-\delta}{\gamma \delta} L_x)$ , solved for,  $l_n$  and  $L_q$  can be recovered using equations 4.74, 4.75, 4.76, and 4.77.

$$\hat{Y} = \left(\frac{1}{\gamma} - 1\right) \mu_n L_x - \tilde{\rho} \quad (4.80)$$

$$= \frac{\delta(1-\gamma) \mu_n}{\gamma} l_n + \frac{\nu \delta \theta \mu_q}{\gamma} \left(L - \frac{\gamma \delta + 1 - \delta}{\gamma \delta} L_x - l_n\right) \quad (4.81)$$

$$l_n = \frac{(1-\gamma) \mu_n + \nu \delta \theta \mu_q (\gamma \delta + 1 - \delta)}{\gamma \delta^2 ((1-\gamma) \mu_n - \nu \delta \theta \mu_q)} L_x - \frac{\nu \delta \theta \mu_q L + \tilde{\rho} \gamma}{(1-\gamma) \mu_n \delta - \nu \delta \theta \mu_q} \quad (4.82)$$

The solution for  $L_x$  has already been found and can be inserted in the expression above. The number of quality-researchers is simply given by  $L - \frac{\gamma \delta + 1 - \delta}{\gamma \delta} L_x - l_n$ .

The growth rate of the economy can be found by inserting the solution for  $L_x$  in equation 4.77.

$$\hat{Y} = \left(\frac{\nu b^2 \theta \mu_q}{(1-\gamma) \mu_n}\right)^{\frac{\delta(\gamma-1)}{\gamma(\delta-1)}} \frac{\delta(1-\gamma) \mu_n}{1-\delta} - \tilde{\rho} \quad (4.83)$$

The growth rates of quality and variety can be retrieved by inserting the solutions for  $L_q$  and  $l_n$  into equations 4.75 and 4.76, respectively.

Of course, endogenous growth is also possible when spillovers are strong enough in only one of the sectors. For example, the specification  $\hat{n} = l_n \mu_n$  with  $\hat{q} = 0$  has been treated by, amongst others, Romer (1990) and the specification  $\hat{q} = \theta \mu_q l_q$  with  $\hat{n} = 0$  has been used by, for example, Grossman and Helpman (1991c). These models can be solved in a way analogous to the procedure applied above.

Another possibility is to generalize the spillovers leading to endogenous growth. Consider the situation where  $\varphi_q = \varphi_n = 0$  and  $K = 1$ . If the growth rates of  $q$  and  $n$  are to be independent of  $n$  and  $q$  along a sustainable growth path, both of these equations should be homogenous of degree zero in  $n$  and  $q$ . This can only be the case if  $\varepsilon_{nq} = 2 - \varepsilon_{qq}$  and  $\varepsilon_{qn} = 1 - \varepsilon_{nn}$ , which yields the following R&D equations:

$$\hat{q} = \theta \mu_q n^{1-\varepsilon_{qq}} q^{\varepsilon_{qq}-1} L_q \quad (4.84)$$

$$\hat{n} = \mu_n n^{\varepsilon_{nn}-1} q^{1-\varepsilon_{nn}} l_n \quad (4.85)$$

Following the same procedure as above, we can equate both versions of the Ramsey rule to get an expression for  $L_x$ .

$$L_x = \left( \frac{\nu b^2 \theta \mu_q}{(1-\gamma) \mu_n} \right)^{\frac{\delta(\gamma-1)}{\gamma(\delta-1)}} \frac{\gamma \delta}{1-\delta} \left( \frac{n}{q} \right)^{\frac{(2-\varepsilon_{qq}-\varepsilon_{nn})\delta(\gamma-1)}{\gamma(\delta-1)}} \quad (4.86)$$

Unfortunately, this exercise does not yield an immediate solution for  $L_x$  as it is a function of  $n/q$  for the general case of  $\varepsilon_{qq} \neq \varepsilon_{nn}$ . In this general case,  $L_x$  can only be constant if  $g_q = g_n$ . That this condition is indeed satisfied can be checked by looking at the growth rate of  $\hat{q}/\hat{n}$ .

$$\frac{d \ln(\hat{q}/\hat{n})}{dt} = (\varepsilon_{qq} + \varepsilon_{nn} - 2) (\hat{q} - \hat{n}) \quad (4.87)$$

As long as  $\varepsilon_{qq} + \varepsilon_{nn} < 2$ , the relative difference between  $\hat{q}$  and  $\hat{n}$  will converge to zero as time proceeds.

The equilibrium ratio  $q/n$  can simply be found by equating  $\hat{q}$  with  $\hat{n}$ .

$$\frac{q}{n} = \left( \frac{\mu_n l_n}{\theta \mu_q L_q} \right)^{\frac{1}{\varepsilon_{qq} + \varepsilon_{nn} - 2}} \quad (4.88)$$

Insert this into equation 4.86 to get

$$L_x = \left( \frac{\nu b^2 \theta \mu_q}{(1-\gamma) \mu_n} \right)^{\frac{\delta(\gamma-1)}{\gamma(\delta-1)}} \frac{\gamma \delta}{1-\delta} \left( \frac{\mu_n l_n}{\theta \mu_q L_q} \right)^{\frac{\delta(\gamma-1)}{\gamma(\delta-1)}}. \quad (4.89)$$

Solutions for  $L_x$ ,  $l_n$ , and  $L_q$  can be obtained using  $L_q = L - \frac{\gamma\delta+1-\delta}{\gamma\delta} L_x - l_n$  and 4.82. The solutions found for  $L_x$  and  $q/n$  can be inserted into 4.85 in order to get the sustainable growth rate of the economy. From equation 4.68 follows that  $\pi/v_n$  is constant, and thus  $\tilde{\rho} = \rho$ .

Summarizing, four variants of endogenous growth can be distinguished. The first two are the familiar variants with either variety or quality growth. The third and fourth variants yield endogenous growth with simultaneous quality and variety growth. The third variant discussed above differs from the fourth in that both kinds of research activities effectively operate independently from each other. In the fourth, generalized, variant the productivity of quality researchers depends very much on variety research and vice versa.

## 4.7 Semi-endogenous growth

As has been mentioned in the introduction to this chapter, models with endogenous economic growth were criticized by Jones (1995a, 1995b) for their prediction that population growth would cause economic growth to accelerate. Jones (1995a) showed that when spillovers are less strong, a constant rate of economic growth is compatible with positive and constant population growth. Li (2000) generalized the result by Jones to include simultaneous quality and variety research. The solutions by Jones and Li both are characterized by the outcome that economic growth ultimately depends on the growth rate of the population. Without population growth, no sustained economic growth is possible in their models.

In the previous section, we saw that endogenous growth arises when  $\hat{q}$  and  $\hat{n}$  are ultimately independent of  $q$  and  $n$ . It was shown that this is the case when  $\hat{q}$  and  $\hat{n}$  are homogenous of degree zero in  $q$  and  $n$ . We will now consider the case where  $\hat{q}$  and  $\hat{n}$  are decreasing in  $q$  and  $n$  for a constant number of researchers, such that, for a constant size of the population, the growth of the economy will diminish over time. In other words, we will now consider the case where the degree of homogeneity of  $\hat{q}$  and  $\hat{n}$  in  $q$  and  $n$  is less than zero. The restrictions that bring about a negative degree of homogeneity can be deduced directly from equations 4.35 and 4.36, which are repeated below for  $K = 1$  and  $L_q = nl_q$ .

$$\hat{q} = L_q^{1-\varphi_q} \theta \mu_q n^{\varepsilon_{nq} + \varphi_q - 1} q^{\varepsilon_{qq} - 1} \quad (4.90)$$

$$\hat{n} = l_n^{1-\varphi_n} \mu_n n^{\varepsilon_{nn} - 1} q^{\varepsilon_{qn}} \quad (4.91)$$

The conditions for a negative degree of homogeneity are  $\varepsilon_{nq} < 2 - \varepsilon_{qq} - \varphi_q$  for quality research and  $\varepsilon_{qn} < 1 - \varepsilon_{nn}$  for variety research.

A negative degree of homogeneity implies that sustained quality or variety growth is not feasible if the population is constant. However, with a growing population, sustained economic growth is possible as can be demonstrated by assuming constant growth rates for  $q$  and  $n$ . Having  $\hat{q} = g_q$  and  $\hat{n} = g_n$ , the system of research equations turns into an ordinary system of two linear equations.

$$0 = (1 - \varphi_q) g_L + (\varepsilon_{nq} + \varphi_q - 1) g_n + (\varepsilon_{qq} - 1) g_q \quad (4.92)$$

$$0 = (1 - \varphi_n) g_L + (\varepsilon_{nn} - 1) g_n + \varepsilon_{qn} g_q \quad (4.93)$$

The solutions to this system are

$$g_q = \frac{(1 - \varphi_q)(\varepsilon_{nq} + \varphi_q - \varepsilon_{nn}) + (\varepsilon_{nq} + \varphi_q - 1)(\varphi_q - \varphi_n)}{(1 - \varepsilon_{qq})(\varepsilon_{nq} + \varphi_q - \varepsilon_{nn}) - (\varepsilon_{nq} + \varphi_q - 1)(\varepsilon_{qn} - \varepsilon_{qq} + 1)} g_L \quad (4.94)$$

$$g_n = \frac{(\varepsilon_{qn} - \varepsilon_{qq} + 1)(1 - \varphi_q) + (\varphi_q - \varphi_n)(1 - \varepsilon_{qq})}{(1 - \varepsilon_{qq})(\varepsilon_{nq} + \varphi_q - \varepsilon_{nn}) - (\varepsilon_{nq} + \varphi_q - 1)(\varepsilon_{qn} - \varepsilon_{qq} + 1)} g_L \quad (4.95)$$

These solutions are not defined if parameter settings cause a division by zero. This happens when the following condition is satisfied:

$$((\varepsilon_{qq} = 1) \vee (\varepsilon_{nn} = \varepsilon_{nn} - \varphi_q)) \wedge ((\varepsilon_{nq} = 1 - \varphi_q) \vee (\varepsilon_{qn} = \varepsilon_{qq} - 1)). \quad (4.96)$$

Table 4.3 contains the four pairs of R&D equations for which the solutions are not defined together with the corresponding parameter settings. Type (i) belongs to the class of endogenous growth models as  $\hat{q}$  is homogenous of degree zero in  $n$  and  $q$ . The steady state growth rate of product variety is zero. Types (ii) and (iii) will be treated in the next section on neo-endogenous growth. A brief discussion of type (iv) can be found in the appendix.

**Table 4.3.** Four special cases of growth models

	Equations	Conditions
(i)	$\hat{q} = L_q^{1-\varphi_q} \theta \mu_q$ $\hat{n} = l_n^{1-\varphi_n} \mu_n n^{\varepsilon_{nn}-1} q^{\varepsilon_{qn}}$	$\varepsilon_{qq} = 1$ $\varepsilon_{nq} = 1 - \varphi_q$
(ii)	$\hat{q} = L_q^{1-\varphi_q} \theta \mu_q n^{\varepsilon_{nq}+\varphi_q-1}$ $\hat{n} = l_n^{1-\varphi_n} \mu_n n^{\varepsilon_{nn}-1}$	$\varepsilon_{qq} = 1$ $\varepsilon_{qn} = \varepsilon_{qq} - 1$
(iii)	$\hat{q} = L_q^{1-\varphi_q} \theta \mu_q q^{\varepsilon_{qq}-1}$ $\hat{n} = l_n^{1-\varphi_n} \mu_n q^{\varepsilon_{qn}}$	$\varepsilon_{nq} = \varepsilon_{nn} - \varphi_q$ $\varepsilon_{nq} = 1 - \varphi_q$
(iv)	$\hat{q} = L_q^{1-\varphi_q} \theta \mu_q n^{\varepsilon_{nn}-1} q^{\varepsilon_{qq}-1}$ $\hat{n} = l_n^{1-\varphi_n} \mu_n n^{\varepsilon_{nn}-1} q^{\varepsilon_{qn}-1}$	$\varepsilon_{nq} = \varepsilon_{nn} - \varphi_q$ $\varepsilon_{qn} = \varepsilon_{qq} - 1$

Returning to the cases that do yield valid solutions for  $g_q$  and  $g_n$ , the growth rate of final production can be found by means of equation 4.66.

$$g_y = \left( \frac{1}{\gamma} - 1 \right) \delta g_n + \nu \delta g_q + g_L \quad (4.97)$$

Here,  $g_n$  and  $g_q$  depend on  $g_L$  by equations 4.95 and 4.94. The steady state growth rate of the economy depends on the growth rate of the workforce and on parameters that cannot be changed by taxes or subsidies. In this sense, economic growth is 'exogenous' rather than 'endogenous'. However, economic growth is not exogenous in the sense that it crucially depends on endogenous research activities, hence the term 'semi-endogenous' growth.

The allocation of the workforce over production, variety research, and quality research can be derived by equating the expression above with the Ramsey rules 4.72 and 4.73. After substitution of  $\hat{n}$ ,  $\hat{q}$ , and  $L_x$  with  $g_n$ ,  $g_q$ , and  $\frac{\gamma\delta}{\gamma\delta+1-\delta} (L - L_q - l_n)$ , respectively, two expressions emerge that have  $l_n$  and  $L_q$  as unknowns. These expressions can be used to obtain (numerical) solutions for  $l_n$  and  $L_q$ .

A remarkable feature of semi-endogenous growth models is that they, contrary to endogenous growth models, can be solved by using only the R&D equations. The reason why it is easier to obtain a solution for the growth rates in semi-endogenous growth models is that the precise allocation of labor does not affect economic growth – as long as employment in both quality and variety research is positive. In endogenous growth models, the allocation of labor was of direct importance for the rate of economic growth and, consequently, the exact allocation had to be found in order to solve the model.

## 4.8 Neo-endogenous growth

Four special cases of growth models were identified in the previous section. Of these four models, the second arguably has the most interesting properties. The R&D equations belonging to model (ii) are repeated below for the case  $\varphi_q = \varphi_n = 0$ .

$$\hat{q} = L_q \theta \mu_q n^{\varepsilon_{nq}-1} \quad (4.98)$$

$$\hat{n} = l_n \mu_n n^{\varepsilon_{nn}-1} \quad (4.99)$$

The peculiar feature of this pair of equations is that they give rise to neo-endogenous growth. At first sight, both equations seem to constitute a clear-cut case of semi-endogenous growth as they have a negative degree of homogeneity in  $n$  while  $q$  does not appear in the equations. Interestingly, this does not turn out to be correct. In an economy with a fixed population size, no variety research will be performed in the steady state – just like in the semi-endogenous case. Without variety research, the number of products is constant and, consequently, so is the productivity of quality researchers. Therefore, as soon as the steady state has been reached and  $n$  is constant, the behavior of the model changes into that of an endogenous growth model with a constant growth rate of quality. The absence of variety expansion causes the ‘dilution effect’ to disappear: no additional quality researchers are needed to staff the research departments of new firms. Without the dilution effect, spillovers are strong enough to ensure sustained quality growth with a constant number of researchers.

Neo-endogenous growth models can be analyzed in two ways. With the first approach the comparative static effects of a change in the size of the population on the rate of economic growth are treated. Population growth is assumed to be absent. The second approach does allow for a growing population, but involves additional restrictions on parameters. Below, the model will be analyzed in both ways. The comparative static method will be treated first, followed by the dynamic approach.

#### 4.8.1 Comparative static analysis

In the steady state without population growth no researchers are employed to develop new products. In this situation all labor will be allocated either to production or to quality research, implying  $L_q = L - \frac{\gamma\delta+1-\delta}{\gamma\delta}L_x$ . In order to solve for the number of products in the steady state, we will study the moment at which the last products are being invented. At this time the rates of return for variety and quality research are still equal, while the number of researchers is approximately zero. By equating the rates of return using equations 4.68, 4.71, 4.98, and 4.99 an expression for  $L_x$  in terms of  $n$  can be found.

$$L_x = \left( \frac{\nu b^2 \theta \mu_q}{(1-\gamma) \mu_n} \right)^{\frac{\delta(\gamma-1)}{\gamma(\delta-1)}} \frac{\gamma\delta}{1-\delta} n^{\frac{(\epsilon_{nq}-\epsilon_{nn})\delta(\gamma-1)}{\gamma(\delta-1)}} \quad (4.100)$$

A second expression for  $L_x$  can be derived by equating 4.74 and 4.72, substituting for  $\hat{q}$ ,  $L_q$ , and  $\hat{n}$ , and letting  $l_n$  approach zero.

$$g_Y = \nu\delta\theta\mu_q n^{\epsilon_{nq}-1} \left( L - \frac{\gamma\delta+1-\delta}{\gamma\delta} L_x \right) = \frac{1-\gamma}{\gamma} L_x \mu_n n^{\epsilon_{nn}-1} - \tilde{\rho} \quad (4.101)$$

$$L_x = \frac{\nu\delta\theta\mu_q n^{\epsilon_{nq}-1} L + \tilde{\rho}}{\frac{1-\gamma}{\gamma} \mu_n n^{\epsilon_{nn}-1} + \nu\delta\theta\mu_q n^{\epsilon_{nq}-1} \frac{\gamma\delta+1-\delta}{\gamma\delta}} \quad (4.102)$$

Again,  $L_x$  is found as a function of  $n$ . The two expressions for  $L_x$  can be used to find the steady state value of  $n$ . In general, no analytical solution is feasible for  $n$ , so a solution has to be obtained using numerical methods. Once the steady state value of  $n$  has been computed, the steady state value of  $L_x$  is easily found, and so is the rate of economic growth. From equation 4.68 follows that  $\pi/v_n$  is constant in the steady state, and thus  $\tilde{\rho} = \rho$ .

According to equation 4.73, the steady state values of  $\hat{y}$  and  $\hat{q}$  can only be constant if  $L_x^{(\gamma-\delta)/[\delta(\gamma-1)]} L_q^{-1}$  is constant as well. This implies that the elasticity of  $L_q$  with respect to  $L_x$  has to satisfy

$$\frac{\Delta \ln L_q}{\Delta \ln L_x} = \frac{\gamma - \delta}{\delta(\gamma - 1)}, \quad (4.103)$$

otherwise the the size of the population will affect the rate of economic growth. From equation 4.100 the elasticity of  $L_x$  with respect to  $n$  can be derived, while the elasticity of  $L_q$  with respect to  $n$  follows from the R&D equation for quality.

$$\frac{\Delta \ln L_x}{\Delta \ln n} = \frac{(\varepsilon_{nq} - \varepsilon_{nn}) \delta (\gamma - 1)}{\gamma (\delta - 1)} \quad (4.104)$$

$$\frac{\Delta \ln L_q}{\Delta \ln n} = 1 - \varepsilon_{nq} \quad (4.105)$$

Divide both elasticities in order to get another expression for the elasticity of  $L_q$  with respect to  $L_x$ .

$$\frac{\Delta \ln L_q}{\Delta \ln L_x} = \frac{\gamma (\delta - 1) (1 - \varepsilon_{nq})}{(\varepsilon_{nq} - \varepsilon_{nn}) \delta (\gamma - 1)} \quad (4.106)$$

Equate 4.103 with 4.106 to see that the size of the population will not affect the growth rate of the economy if the following condition is satisfied:

$$\varepsilon_{nq} = \frac{\gamma (\delta - 1)}{\delta (\gamma - 1)} + \frac{\gamma - \delta}{\delta (\gamma - 1)} \varepsilon_{nn} \quad (4.107)$$

Although neo-endogenous growth is feasible, it requires that the condition in 4.107 holds while  $\varphi_q = \varphi_n = 0$ . If one of these restrictions is violated, such that larger economies grow at a slower pace than smaller economies, growth is semi-endogenous; if the reverse occurs, growth is endogenous.

## 4.8.2 Dynamic analysis

The comparative static analysis presented above showed that neo-endogenous growth is only feasible if the elasticity of  $L_q$  with respect to  $L_x$  satisfies the condition in 4.103. According to this condition,  $L_q$  and  $L_x$  will only grow at the same rate for  $\delta = 1$  – that is, if growth is to be neo-endogenous. With  $\delta = 1$ , the condition in 4.107 implies a further restriction on the parameters:  $\varepsilon_{nq} = \varepsilon_{nn} \equiv \varepsilon_n$ . Applying the latter restriction to the R&D equations returns

$$\hat{q} = \theta \mu_q L_q n^{\varepsilon_n - 1} \quad (4.108)$$

$$\hat{n} = \mu_n l_n n^{\varepsilon_n - 1}. \quad (4.109)$$

If both quality and variety research are taking place, the rates of returns of the two activities have to be equal. This results in an additional restriction on the parameters.

$$\left( \frac{1}{\gamma} - 1 \right) \frac{L_x}{l_n} \hat{n} = \frac{\nu b^2}{\gamma} \frac{L_x}{L_q} \hat{q} \Rightarrow \frac{\mu_n}{\theta \mu_q} = \frac{\nu b^2}{1 - \gamma} \quad (4.110)$$

If the condition above is satisfied, the rates of return for quality and variety research are always equal. This phenomenon has profound consequences for the analysis of the model. For the special case  $\delta = 1$ , there is no mechanism that

uniquely determines the allocation of labor between the two research activities. When the number of researchers working on quality improvements or variety expansion is indeterminate, then the same applies to the growth rate of the economy. Although we cannot find a unique solution for the growth rate, it is still worthwhile to proceed with the analysis. For this purpose, we assume that  $L_x/L_q$  equals the constant  $\Lambda$ .

When the economy is in the steady state, the growth rate of the number of products is a function of the growth rate of population,  $g_n = (1 - \varepsilon_n)g_L$ . The steady state growth of output is therefore given by

$$g_y = \frac{1 - \varepsilon_n}{\gamma} g_L + \nu g_q \quad (4.111)$$

What remains to be found, is the growth rate of the quality level. This can be done by combining the expression above with the Ramsey rule (equation 4.73).

$$g_q = \frac{\gamma - 1 + \varepsilon_n}{\gamma\nu - \nu b^2\Lambda} g_L - \frac{\gamma\rho}{\gamma\nu - \nu b^2\Lambda} \quad (4.112)$$

Here,  $\bar{\rho}$  has been replaced with  $\rho$  as  $\pi/v_n$  is constant in the steady state (equation 4.68).

Enter the found expression for  $g_q$  into equation 4.111 to get the final result.

$$g_y = \left( \frac{1 - \varepsilon_n}{\gamma} + \frac{\gamma - 1 + \varepsilon_n}{\gamma - b^2\Lambda} \right) g_L - \frac{\gamma\rho}{\gamma - b^2\Lambda} \quad (4.113)$$

The remarkable feature of the solution for steady state economic growth is that it only partly depends on  $g_L$ . Population growth has a positive effect on economic growth, but economies with a constant population are also capable of growing – as long as  $\gamma < b^2\Lambda$ . In other words, growth is a mixture of endogenous and semi-endogenous growth.

### 4.8.3 Network-confined spillovers

Peretto and Smulders (2002) have presented a modified version of the neo-endogenous growth model discussed above. In their version knowledge spillovers are confined to a network. Only researchers working in the same network can benefit from each other's ideas. They assume that each network has the same size and they impose the following restrictions on the parameters:  $\varphi_q = \varphi_n = 0$ ,  $\eta_q = \eta_n = 1$ ,  $\zeta_q = 1 - \varepsilon_q$ , and  $\zeta_n = -\varepsilon_n$ . Consequently, the arrival rates in 4.31 can be written as

$$\phi_q = \mu_q \left( \frac{n}{K} \right)^{\varepsilon_{nq}} q \quad (4.114)$$

$$\phi_n = \mu_n \left( \frac{n}{K} \right)^{\varepsilon_{nn}}, \quad (4.115)$$

where the number of networks evolves according to  $K = K_0 + \kappa(n - n_0)$ .

The R&D equations corresponding to the arrival rates are

$$\hat{q} = \mu_q \frac{L_q}{n} \left( \frac{n}{K} \right)^{\varepsilon_{nq}} \quad (4.116)$$

$$\hat{n} = \mu_n \frac{l_n}{n} \left( \frac{n}{K} \right)^{\varepsilon_{nn}}. \quad (4.117)$$

Apart from the term involving  $n/K$ , this specification is identical to the specification used with ordinary neo-endogenous growth models. For very large economies

the ratio  $n/K$  converges to  $\kappa^{-1}$ , causing the model to approach the standard model of neo-endogenous growth. For 'small' economies, the Peretto-Smulders approach causes the growth rate of the economy to be affected by its size. This 'scale effect' may be positive or negative, depending on the initial values of  $K_0$  and  $n_0$ . The sign of the scale effect equals the growth rate of  $n/K$ .

$$\hat{n} - \hat{K} = \frac{K_0 - \kappa n_0}{n(K_0 + \kappa(n - n_0))} \quad (4.118)$$

The scale effect is positive if the number of firms per network is smaller than  $\kappa^{-1}$ ; if it is larger, the scale effect is negative.

#### 4.8.4 Type (iii) spillovers

Type (iii) spillovers may give rise to either endogenous or neo-endogenous growth. For  $\varepsilon_{qn} \geq 0$ , spillovers in variety research are strong enough to generate endogenous growth. When  $\varepsilon_{qn} < 0$ , the structure of spillovers is identical to that of a type (ii) model, but then with  $n$  and  $q$  interchanged. Technically, this modification is of little significance and the model can be solved in much of the same way as the standard neo-endogenous growth model.

However, the interpretation of neo-endogenous growth arising from type (iii) spillovers is quite different. A higher average quality level makes it more difficult to invent a new product as  $\varepsilon_{qn} < 0$ . Consequently, as long as  $\hat{q} > 0$  an increasing number of variety researchers is needed in order to keep  $\hat{n}$  constant. Because the growth rate of quality is negatively affected by its level, the steady state quality level will be constant if population growth is absent. With the economy in the steady state, the negative spillover from quality to variety comes to a hold and sustained variety growth becomes feasible.

The plausibility of this type of neo-endogenous growth is open to debate. On the one hand, there is no reason to assume that new products are more difficult to discover if the quality of existing products is higher. The reverse seems to be more likely. On the other hand, it does seem sensible to assume that more researchers are needed if a new product has to have a higher initial quality level.

### 4.9 Conclusion

The early endogenous growth models have been criticized for their prediction that population growth causes the rate of economic growth to increase without bound. The semi-endogenous growth models developed in response to these early models did not have this 'scale effect'. An important feature of neo-endogenous growth is that it crucially depends on the rate of population growth. Starting with Young (1998), several authors have presented models that generated economic growth in economies with a fixed population, but that did not exhibit explosive growth when population growth is positive.

A theoretical critique on models characterized by (neo-) endogenous growth concerns the generality of the specifications used for knowledge spillovers. Jones (1995a) and later Li (2000) have shown that (neo-) endogenous growth will only occur for specific parameter settings – so-called 'knife-edge' conditions. In contrast, semi-endogenous growth is feasible over a range of parameter settings. The lack of generality of (neo-) endogenous growth models suggests that economic growth without population growth can only occur under rather exotic circumstances.



What do models of economic growth with intertemporal knowledge spillovers tell us about the relation between product variety and the size of the population? Endogenous growth and neo-endogenous growth with type (iii) spillovers allow for a growing number of product types when the size of the population is fixed. With positive population growth endogenous growth implies explosive growth in the number of products, while the growth rate of product variety is stable for neo-endogenous growth. A stricter relation between product variety and population size is predicted by models with semi-endogenous growth and standard (type (ii)) neo-endogenous growth. In these models growth in the number of products can only be sustained when the population is growing.

Above, only the effects on product variety caused by population growth and knowledge spillovers have been taken into account. Chapters 7 and 8 will focus on other determinants of product variety. In chapter 7, a model will be presented that allows for changes in the level and distribution of education. Chapter 8 will provide an analysis of how economic integration affects the number of product types that are produced and consumed in a country.

#### 4.A Type (iv) spillovers

The type (iv) R&D-equations are both homogenous of a degree smaller than zero and therefore models with this type of spillovers are characterized by semi-endogenous growth. What makes the type (iv) model deviate from the class of strong semi-endogenous growth models is that the R&D equations alone do not yield a solution for the steady state growth rates.

$$0 = (1 - \varphi_q) g_L + (\varepsilon_{nn} - 1) g_n + (\varepsilon_{qq} - 1) g_q \quad (4.119)$$

$$0 = (1 - \varphi_n) g_L + (\varepsilon_{nn} - 1) g_n + (\varepsilon_{qq} - 1) g_q \quad (4.120)$$

A first observation is that steady state growth is possible only when  $\varphi_q = \varphi_n = \varphi$ .<sup>12</sup> In this case, however, both equations are identical, which means that these equations do not provide enough information to be able to solve for  $g_q$  and  $g_n$  in terms of the parameters of the model.

Additional information can be derived from the fact that the rate of return for quality research has to equal the rate of return for variety research,  $\pi_q \hat{q}/v_q = \pi/v_n$ . Substitution of  $\hat{q}$  and  $\hat{n}$  in equations 4.68 and 4.71 and equating both rates of return yields  $L_x$  as a function of  $L_q/l_n$ .

$$L_x = \left( \frac{(1 - \gamma) \mu_n}{\nu b^2 \theta \mu_q} \right)^{\frac{\delta(\gamma-1)}{\gamma(1-\delta)}} \frac{\gamma \delta}{1 - \delta} \left( \frac{L_q}{l_n} \right)^{\frac{\varphi \delta(\gamma-1)}{\gamma(1-\delta)}} \quad (4.121)$$

Remarkably,  $L_x$  does not depend on the growth rate of the population. This implies that when the population grows, employment in the intermediate and final goods sectors remains constant and thus that  $L_q$  and  $l_n$  grow a rate that is (slightly) higher than that of the total population. The constancy of  $L_x$  has far-reaching consequences. In a growing population, the share of the population employed in ordinary production activities declines over time, ultimately leading to overemployment in research activities.

It can be shown that sustainable growth is not feasible when the population is growing. The rate of return to (variety) research should be constant over time if per capita income is to grow at a steady rate (see 4.68).

$$g_{\pi/v_n} = -\varphi g_L + (\varepsilon_{nn} - 1) g_n + (\varepsilon_{qq} - 1) g_q = 0 \quad (4.122)$$

It is immediately clear that this condition is in contradiction with the steady state growth conditions that have been derived from the R&D equations above. The fixed amount of labor employed for the production of final and intermediate goods drives down the rate of return for research activities whenever the population is growing. This implies that no research will be done in the steady state and that there will be no growth in per capita consumption. In the absence of population growth, steady state growth is also not feasible as the degree of homogeneity of the R&D equations is negative. The constancy of  $L_x$ , however, does ensure that research activities do not vanish over time as long as  $L_q/l_n$  is constant. The precise allocation of labor remains indeterminate. For the special case  $\varphi = 0$ , it is possible to obtain a solution for  $L_x$ . This solution follows directly from the expression for  $L_x$  derived above.

## Notes

1. A (critical) examination of this and other empirical facts that have, or should have, shaped the theoretical literature on economic growth is provided in chapter 3.
2. Some important effects of randomness in models of economic growth are discussed in chapter 9.
3. In case  $\delta = 1$  and  $y$  is a consumption index, the optimization problem has to be formulated differently as consumers do not maximize profits but utility. The alternative formulation is

$$\mathcal{L}(j) = \max_{x(ij)} \left\{ y(\mathbf{x}(j)) + \lambda_y \left( B(j) - \int_N p(i) x(ij) di \right) \right\},$$

where  $B$  is the budget constraint for consumption goods and  $\lambda_y = dy/dB = 1$  as  $y$  is the numeraire. The first order conditions are identical to the profit maximization problem.

4. A generalization that would involve both a variable  $y$  and  $n_y$  would require that each intermediate firm does not only choose the amount to supply to each firm that produces final output, but also chooses the number of firms to supply.
5. The assumption that researchers work independently, even when they are part of the same project, facilitates the analysis of the model as it neutralizes the stochastic nature of inventions. Many of the authors that do not make this assumption fail to treat randomness in a satisfactory way. Relatively elegant treatments can be found in Dinopoulos and Thompson (1998) and Thompson (2001).
6. The stock of previous inventions is occasionally taken to affect the size of innovations ( $\theta$ ), rather than their arrival rate ( $\phi$ ). In the current setup where uncertainty is neutralized both approaches are equivalent.
7.  $|Q(i)|$  is the size of set  $Q(i)$ .
8. Alternatively, a transversality condition could be formulated that prevents unsustainable debts. As consumer heterogeneity is strictly forbidden here, no net debt is possible. In the current context, a transversality condition would be too weak a requirement.
9. If a firm would issue new shares without using them productively, the new funds would merely inflate dividends. The symmetry of households implies that the households that buy the new shares will directly receive the same amount in return as dividends.

10. For an excellent discussion of dynamic optimization see Kamien and Schwartz (1991), especially pages 124-127.
11. A more general specification of utility function would be the CIES function:  $u(c) = \frac{c^{1-\theta}-1}{1-\theta}$ . The log-utility function corresponds to  $\theta \rightarrow 1$ .
12. If  $\varphi_q$  and  $\varphi_n$  are allowed to be different, the research activity with the highest  $\varphi$  will asymptotically vanish, such that only one R&D sector will remain active.

## A measure of product variety

The concept of effective product variety was already introduced in chapter 1 – albeit in a rather informal way. In this chapter, the antilog of Shannon’s entropy is proposed as a formal measure of effective variety. Three properties of this ‘variety index’ are of particular relevance. First, for symmetric product types the index is equal to the number of product types. Second, disaggregation of the underlying product set always leads to an increase in measured product variety. Third, the introduction or disappearance of a marginal type does not cause a discrete change in the variety index. The proposed variety index will be used in chapters 6, 7, 9, and 10.

### 5.1 Introduction

Suppose that a person is asked to abstain from consuming one type of product. In the highly unlikely case that this person is indifferent about which type will be removed from his shopping cart, we can say that all types of products are equivalent to this consumer. When product types are equivalent, consumer welfare simply depends on the number of products and the total quantity of goods consumed.<sup>1</sup> It is this special case of equivalent types that we have encountered in the previous chapter.

The simple relation between the number of products and the standard of living breaks down if a new type of good is introduced that is not equivalent to the other types. If the consumer considers the new type to be inferior compared to the other types, then the welfare of the consumer will not rise as much as we would have expected if we would have only taken into account the number of product types available to him. For the consumer the standard of living hardly changes as he regards the new product to be of little use.

This small thought-experiment suggests that an ideal measure of product variety should take into account the importance of each product. Additionally, it seems natural to let such a measure be equal to the number of product types if all types are equivalent. In other words, an ideal measure of product variety should be something like a ‘weighted number of products’. It is argued in this chapter that the antilog of Shannon’s entropy has a number of properties that makes it a suitable candidate for the ideal indicator of product variety.

The concept of entropy originates from the work of Boltzmann in thermodynamics.<sup>2</sup> Later on, Shannon (1948) developed a measure for the amount of

information that is lost when communicating over a noisy channel. His friend Von Neumann gave him two reasons to call his measure ‘entropy’:

First, the expression [5.1] is the same as the expression for entropy in thermodynamics and as such you should not use two different names for the same mathematical expression, and second, and more importantly, entropy, in spite of hundred years of history, is not very well understood yet and so as such you will win every time you use entropy in an argument! (Kapur and Kesavan 1992, p. 8)

Entropy was introduced to economics by Henry Theil (1965, 1967), who borrowed it from Shannon. Theil (1967, pp. 91, 290) has suggested the use of entropy as a measure of income inequality and industry-concentration. More recently, Alexander (1997) and Frenken, Saviotti, and Trommetter (1999) have applied Shannon’s entropy as an indicator of product variety. In these studies, product types have been weighted according to the frequency of their occurrence in a dataset.

Section 5.2 introduces the properties of Shannon’s entropy and shows that they are preserved for a class of weights that includes, but is not limited to, frequencies. This property allows us to weigh product types by, for example, their value. Section 5.3 compares Shannon’s entropy with other measures of product variety. Concluding remarks can be found in section 5.4.

## 5.2 Entropy and product variety

Suppose that we want to measure the variety of a set  $N$  that contains  $n$  different types of products. Each product type  $i \in N$  is assigned a weight  $w_i$  that depends on how important the type is compared to the other types as indicated by a measure  $u$ . In particular,  $w_i \equiv u_i / \sum_{j \in N} u_j$ . If not all types have the same weight, then the measurement of variety becomes non-trivial. In this situation, entropy,  $H$ , is a good index of product variety because it possesses the three properties stated below.

**Property 1** *If  $w_i = \frac{1}{n}$  for all  $i$ , then  $H$  is monotonically increasing in  $n$ .*

This property implies equivalence with the most basic variety-index possible, which is the number of types.

**Property 2**  *$H(N) = H(G) + \sum_{j \in G} w_j H(N_j)$  where  $N_j$  is a subset of  $N$  such that  $\cup_{j \in G} N_j = N$  and  $N_j \cap N_k = \emptyset$  for all  $j \neq k$ , and  $G$  is the collection of subsets.*

In words, total variety is the variety of subsets plus the weighted sum of variety within each subset. This property ensures that the total entropy at a higher level of aggregation is lower than the total entropy at a more detailed level of aggregation, even when only one subset is disaggregated.

In order to clarify this point consider the following example. Suppose the set of products ‘fruit’ contains ‘oranges’, ‘apples’, and ‘bananas’. Each type of fruit gets the weight  $1/3$ . The subset ‘apples’ in turn consists of ‘red apples’ and ‘green apples’, each type of apple being weighted equally.

Figure 5.1 displays two ways in which this ‘fruit-basket’ can be represented. In the left panel, the two types of apples are not grouped together, while in the right panel they are. The entropy of the fruit-basket is the same in both panels, but

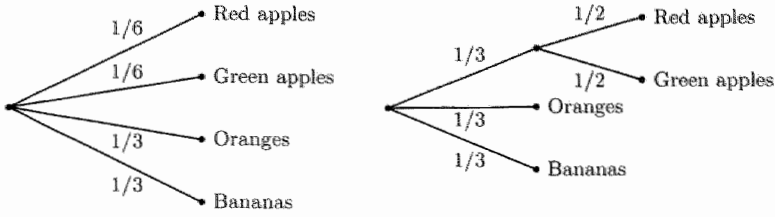


Figure 5.1. Entropy and product classification

each panel suggests a distinct way of computing that entropy. For the left panel, entropy is simply computed as  $H(1/6, 1/6, 1/3, 1/3)$ . For the right panel, entropy equals the entropy of the product groups plus the entropy of the ‘apples-group’ multiplied by the weight of the apples-group:  $H(1/3, 1/3, 1/3) + \frac{1}{3}H(1/2, 1/2)$ . Property 2 ensures that both procedures yield the same outcome.

It is straightforward to demonstrate that entropy increases when a product group is disaggregated. Subtract the entropy of the fruit-basket before disaggregation from the entropy after disaggregation to get the contribution of variety in apples:

$$H(1/6, 1/6, 1/3, 1/3) - H(1/3, 1/3, 1/3) = \frac{1}{3}H(1/2, 1/2).$$

**Property 3**  $H$  is continuous in all  $w_i$ .

This property ensures that when an infinitesimal amount of a product is produced that did not exist previously, the variety index will not make a discrete jump but will change only marginally. By this property, the introduction or disappearance of an unimportant type (i.e. a type with a small weight compared to the other types) does not disrupt the variety index.

**Theorem 3** (Shannon<sup>3</sup>) The only function  $H$  that possesses all three properties stated above is

$$H(N) = -K \sum_{i \in N} w_i \ln w_i \quad (5.1)$$

$$w_i \equiv \frac{u_i}{\sum_{j \in N} u_j} \quad (5.2)$$

where  $K$  is an arbitrary constant, and  $u$  is some given measure along which the elements of  $N$  (the product types) can be compared such that

$$u_N = \sum_{j \in N} u_j. \quad (5.3)$$

**Proof.** Define  $m_i \equiv u_i/\bar{u}$ , where  $\bar{u}$  is an arbitrary constant. By choosing  $\bar{u}$  very small, the  $m_i$  can get arbitrarily close to integer values i.e.  $\max\{m_i - \lfloor m_i \rfloor\}_{i \in N} \simeq 0$ . We will start by showing that  $H(M)$  possesses properties 1 and 2, where  $M$  is a set containing  $m$  elements of size  $\bar{u}$  (i.e.  $u_i = \bar{u} \forall i \in M$ ).

By definition, all elements of  $M$  are equally weighted, such that  $w_i = 1/m \forall i \in M$ . In this symmetric case,  $H$  depends only on  $m$  and the notation can be simplified into  $H(M) = A(m)$ . Suppose that the units are classified into  $s$  broad

subsets and that all these subsets are again subdivided into  $s$  subsets and so on until the number of times sets are subdivided is  $p$ . After  $p$  steps, the total number of subsets at the most detailed level of disaggregation will be  $s^p$ .

By property 2 this implies that  $A(s^p)$  should satisfy

$$A(s^p) = A(s) + s \frac{1}{s} A(s) + s^2 \frac{1}{s^2} A(s) + \dots + s^{p-1} \frac{1}{s^{p-1}} A(s) \quad (5.4)$$

$$= pA(s). \quad (5.5)$$

In the same way, if the units are  $q$  times subdivided into  $t$  subsets per set, the total number of subsets at the most detailed level is  $t^q$ . From property 2 follows again that  $A(t^q) = qA(t)$ .

For any value of  $q$ ,  $s$ , and  $t$ , a  $p$  can be found such that

$$s^p \leq t^q < s^{p+1}. \quad (5.6)$$

After taking logarithms and dividing by  $q \ln s$  we get

$$\frac{p}{q} \leq \frac{\ln t}{\ln s} < \frac{p}{q} + \frac{1}{q}. \quad (5.7)$$

Rearranging yields

$$0 \leq \frac{\ln t}{\ln s} - \frac{p}{q} < \frac{1}{q} \Rightarrow \left| \frac{\ln t}{\ln s} - \frac{p}{q} \right| < \varepsilon. \quad (5.8)$$

As  $q$  can take on any value,  $\varepsilon$  can be taken to be arbitrarily close to zero.<sup>4</sup>

Property 1 implies that

$$A(s^p) \leq A(t^q) < A(s^{p+1}) \quad (5.9)$$

$$pA(s) \leq qA(t) < (p+1)A(s) \quad (5.10)$$

$$\frac{p}{q} \leq \frac{A(t)}{A(s)} < \frac{p}{q} + \frac{1}{q} \quad (5.11)$$

$$\left| \frac{p}{q} - \frac{A(t)}{A(s)} \right| < \varepsilon, \quad (5.12)$$

and thus

$$\frac{A(t)}{A(s)} - \frac{p}{q} + \frac{p}{q} - \frac{\ln t}{\ln s} < 2\varepsilon \rightarrow \left| \frac{A(t)}{A(s)} - \frac{\ln t}{\ln s} \right| < 2\varepsilon. \quad (5.13)$$

For  $\varepsilon$  arbitrarily close to zero,  $A(t)$  should be of the form  $K \ln t$ . The constant  $K$  should be positive, otherwise  $H$  would not be increasing in  $q$ . The function  $H(M) = K \ln m$  possesses all three properties.

The elements of  $M$  had equal weights. The form of  $H$  for heterogenous weights still has to be found. Fortunately, this general form of  $H$  can be found by the use of property 2. Allocate the  $m$  units of  $M$  over the product types in the set  $N$ . Property 2 then implies

$$A(m) = H(N) + \sum_N w_i A(m_i) \quad (5.14)$$

$$H(N) = \sum_N w_i (A(m) - A(m_i)). \quad (5.15)$$

Substituting for  $A(m) = K \ln m$  yields

$$H(N) = \sum_N w_i \left( K \ln \frac{m}{m_i} \right). \quad (5.16)$$

By  $w_i = m_i \bar{u} / \left( \sum_{j \in N} m_j \bar{u} \right) = m_i / m$  the expression for  $H(N)$  can be written as

$$H(N) = -K \sum_N w_i \ln w_i. \quad (5.17)$$

That property 3 holds for  $w_i > 0$  is trivial. That  $H$  is continuous in  $w_i$  when  $w_i \rightarrow 0$  is demonstrated below.

$$\lim_{w_i \downarrow 0} H = -K \sum_{j \in N/i} w_j \ln w_j - K \lim_{w_i \downarrow 0} (w_i \ln w_i) \quad (5.18)$$

L'Hôpital's rule implies

$$\lim_{w_i \downarrow 0} (w_i \ln w_i) = \lim_{w_i \downarrow 0} \frac{d(w_i^2 \ln w_i)}{dw_i} \quad (5.19)$$

$$= \lim_{w_i \downarrow 0} (2w_i \ln w_i) + \lim_{w_i \downarrow 0} w_i. \quad (5.20)$$

Subtraction of  $2 \lim_{w_i \downarrow 0} (w_i \ln w_i)$  yields  $-\lim_{w_i \downarrow 0} (w_i \ln w_i) = \lim_{w_i \downarrow 0} w_i = 0$ , which proves continuity of  $H$  in  $w_i$ . ■

It is possible to normalize  $H$  in order to resemble more closely the number of product types as a measure of product variety. If we set  $K = 1$  and take the exponential form of  $H$  we get

$$V(N) = \prod_N w_i^{-w_i}. \quad (5.21)$$

The properties of  $V$  that correspond to the three properties of  $H$  are:

**Property 4** If  $w_i = \frac{1}{n}$  for all  $i$ , then  $V = n$ .

**Property 5**  $V(N) = V(G) \prod_{j \in G} V(N_j)^{w_j}$

**Property 6**  $V$  is continuous in all  $w_i$ .

### 5.3 Discussion

Of course, other measures can be thought of as indicators of product variety. In principle, any inequality measure with a normalized domain between zero and one multiplied by the number of product types is a candidate. For example, the Hirschman-Herfindahl index (HHI) can be used as a variety index.

$$\text{HHI} \equiv \sum_N w_i^2 \quad (5.22)$$

While the HHI is monotonically increasing in  $n$  when  $w_i = 1/n$  (property 1) and continuous in  $w_i$  (property 3), it cannot be disaggregated according to property 2 or 5. In fact, as has been shown above that only Shannon's entropy possesses all three properties.



Because the HHI fails to have property 2, decomposition into the contributions of subsets requires a change in the functional form of the index. This has been shown by Mussard, Seyte, and Terraza (2003). They also show that the same applies to the Gini coefficient. Property 2 ensures that Shannon's entropy can be disaggregated without losing its functional form. It is this property that makes (the antilog of) Shannon's entropy superior to indices like the HHI and the Gini coefficient.

If we actually want to compute the variety index, we need to choose an explicit variable for  $u$ . In the context of measuring product variety, an obvious choice is to let  $u_i$  equal the expenditure on good  $i$ . This approach will be taken in chapters 6, 7, and 9.

Using expenditure to weigh product types is a method that is open to criticism. One could argue that types of goods for which the value traded is the same are all equally unique in their ability of generating output or utility. This line of argument ignores the possible existence of close substitutes. Suppose the value traded of two types of goods,  $A$  and  $B$ , is identical but that there exists a good alternative only for good  $A$ . Now, by way of experiment, remove type  $A$  from the economy. The damage caused by this action will be limited as the substitute will take over the functions of  $A$ , albeit less efficiently. Removal of type  $B$  will have far greater consequences as it cannot be replaced by something else. Type  $B$  is more unique than type  $A$  even though they are being weighted equally.

The failure to explicitly take into account substitutability and uniqueness is not as disadvantageous as it might appear to be. Let us return to the 'fruit-basket' example given above and experiment by removing apples. Suppose red and green apples are very close substitutes. The removal of all green apples then would lead to a reduction in entropy of 0.231 as red apples, oranges and bananas all have a weight of  $1/3$ . If there would be no substitute for green apples, the weights would be  $1/5$  for red apples and  $2/5$  each for oranges and bananas. The corresponding decline in entropy is 0.275, which is larger than 0.231, reflecting that the loss in variety is larger when there are no substitutes. Substitution effects do have an influence on value-weighted entropy, although this influence might not be as large as would be desirable.

Diversity value functions like those suggested by Weitzman (1992, 1998) and Nehring and Puppe (2002) explicitly take the similarity of objects as the foundation for the valuation of diversity. Their approach has a substantial advantage over the index proposed here in that it allows for different degrees of substitutability between products. A disadvantage of these diversity value functions is that quantities of types of objects (products, for example) are deemed to be irrelevant for the value of diversity. From an economic point of view, the scarcity of objects should matter for the value of diversity. What good is the existence of red apples if the worldwide annual production is just ten pieces?

## 5.4 Conclusion

The chapter started with arguing that the number of product types is not an ideal measure of product variety as it does not allow for the weighting of product types. The antilog of Shannon's entropy is proposed as an alternative indicator of product variety.

The antilog of entropy possesses three properties that make it suitable for this purpose. First, for symmetric product types it equals the number of product types. Second, disaggregation of the underlying product set always leads to an

increase in measured product variety. Third, the introduction or disappearance of a marginal type does not cause a discrete change in the variety index. It has been shown that these properties are preserved for weights based on any measure  $u$  for which holds that  $u_N = \sum_{j \in N} u_j$ . This result is a modest generalization of Shannon's second theorem.

Indices like the Hirschman-Herfindahl index or the Gini coefficient are inferior to the variety index that is proposed here as these indices cannot be disaggregated without losing their original functional form. Another advantage of the antilog of entropy is that it has direct relevance to the theory of index numbers – something that will be demonstrated in chapter 6.

## Notes

1. It is implicitly assumed here that all goods have the same quality and price. However, if products differ with respect to their quality, they might still be considered equivalent if the difference in quality is exactly compensated for by a difference in price. Strictly speaking, it is the quantity in 'constant quality units' that matters, not the physical quantity.
2. For a clear discussion of the different forms of entropy and their applications see Kapur and Kesavan (1992).
3. The theorem and its proof are more general than the similar theorem by Shannon (1948, theorem 2) as it shows that the properties stated above hold for a class of weights that includes, but is not limited to, frequencies.
4.  $\bar{u}$  can be chosen to be arbitrarily small, implying that  $m$  can be arbitrarily large.



## Indices without baskets

In this chapter a new set of price and quantity indices is introduced. The price and quantity indices are measured over single sets of products, which makes them suitable for the comparison of sets that do not contain the same types of products. The introduction of new goods and the disappearance of old ones do not lead to biases in these indices. The indices do not rely on assumptions and they unconditionally pass the product test,  $E = PQ$ . The variety index of chapter 5 plays a prominent role in this chapter. The indices can be used to decompose changes in productivity and the cost of living, provided that the underlying production or consumption function is homothetic. Most of the results presented below will not be used in the other chapters, but they will still be of interest as they provide a fairly general background on the relation between welfare and product variety.

### 6.1 Introduction

Variety is the *raison d'être* of indices. Without product variety there would hardly be a need for the various indices devised to measure the quantity and price of sets of goods. Indeed, for a set of homogenous goods a satisfactory quantity index would simply be the number of goods in the set, or some equivalent measure like their joint mass or volume. It is also straightforward to compute a representative price, even if for some reason these goods would be traded against different prices, by dividing total expenditure by the total quantity. This amounts to the same thing as a quantity-weighted arithmetic average of prices:  $P = (\sum p_i x_i) / \sum x_j = \sum p_i (x_i / \sum x_j)$ . ( $p_i$  and  $x_i$  are the price and quantity of a type of good respectively;  $P$  is the price index.)

Drobisch (1871) suggested to use this method also for heterogeneous products. In a reaction to Drobisch, Laspeyres (1871) argued that simply dividing total expenditure through the number of goods does not yield a satisfactory price index when products are heterogeneous.

Die Formel von Drobisch genügt nicht, sie kann ad absurdum führen, denn nach derselben kann sich eine durchschnittliche Preissteigerung oder Preissenkung ergeben, während alle einzelnen Waaren in Preise gleich geblieben sind. (Laspeyres 1871, p. 308)

Laspeyres gives the following example to illustrate his criticism. Suppose the price of good *A* is 1 in both periods and the price of good *B* is 2, also in both

periods. The quantity of *A* traded is 100 in period 0 and 1000 in period 1; the quantities of *B* are 100 and 20, respectively. Drobisch' index would then be 0.68, even though the prices of both goods remained constant.

Laspeyres' objection is not entirely justified. It does seem sensible to allow for changes in an index when the weights change, even when all prices remain constant. The problem with Drobisch' approach is a bit more subtle than Laspeyres suggested. Drobisch' method is unappealing because every good is considered to be equally important. As a consequence, a moderate change in the quantity of a type of good with a small budget share can have a substantial impact on the price index. This drawback can in principle be avoided by weighting the price of a good by, for example, its share in total expenditure. For some reason or another, however, this approach to indices has never been popular.

The primary purpose of an index being the comparison of different sets of goods, Laspeyres (1871) and Paassche (1874) devised price indices that measure the difference between sets (so called bilateral indices), avoiding thereby the construction of indices that are absolute (or unilateral) in the sense that they refer to just a single set. Specifically, what Laspeyres and Paassche did was to compare the actual expenditure on one set with the hypothetical expenditure on the other set as if the quantities of the different types of goods would have been the same as in the first set. This 'relative' approach to the construction of indices has become something like a paradigm as virtually every economic index nowadays measures differences between sets and is not defined over a single set.<sup>1</sup>

Two arbitrary decisions have to be made when computing a Laspeyres or Paassche index. First, a reference set, the set that is used for the computation of the weights, has to be chosen. Second, something should be done with types of goods that are present in one set of goods but not in other. For the first problem alternative indices have been devised such as the Fisher (first proposed by Bowley 1899, p. 641) and Sato-Vartia indices (Sato 1976, Vartia 1976).

The common way of dealing with the second problem is to exclude from the index the types of goods only present in one of the sets as if it is simply a problem of lacking data. The following example illustrates why this procedure can lead to substantial biases.<sup>2</sup> Suppose we want to compute the price index for a set containing two goods, *A* and *B*. *A* is produced in both period 1 and 2 with a constant price *p* and quantity *x* ( $p_A^1 = p_A^2 = x_A^1 = x_A^2 = 1$ ); Good *B* is only produced in period 2: ( $x_B^1 = 0$ ,  $p_B^1 = ?$ , and  $p_B^2 = x_B^2 = 1$ ). Treating the unknown price of *B* in period 1 as missing data, the Laspeyres price index evaluates to

$$P_L = \frac{x_A^1 p_A^2 + x_B^1 p_B^2}{x_A^1 p_A^1 + x_B^1 p_B^1} = \frac{1 + 0}{1 + 0*?} = 1,$$

while the Paassche price index evaluates to

$$P_P = \frac{x_A^2 p_A^2 + x_B^2 p_B^2}{x_A^2 p_A^1 + x_B^2 p_B^1} = \frac{1 + 1}{1 + 1*?} = 2.$$

Clearly, the Paassche index is biased. The Laspeyres index seems to be unbiased, but this just accidental: if we reverse periods 1 and 2, implying that good *B* disappears in the second period, we will see that the Laspeyres index is also biased.

The easiest, but rather crude, way to reduce this bias is to exclude all products for which data are lacking (Diewert 1980, p. 501). A more refined approach involves estimating the 'missing' prices. This approach was first suggested by Hicks (1940, p. 114) who stated that the price of an unsold good should equal its reservation price: the price at which the demand for the goods just equals zero.

Later on, Fisher and Shell (1972) and Diewert (1980) also advocated the use of this method. Recently, Hausman (1997, 2003), Petrin (2002), Goolsbee and Petrin (2001), and Nevo (2003) have estimated reservation prices for some specific new goods. As these studies indicate, it is not that simple to get a reliable estimate of the reservation price of a new good.

Feenstra has suggested a somewhat different approach (Feenstra, Markusen, and Zeile 1992, Feenstra 1994, Feenstra and Markusen 1994, Feenstra and Kee 2004) (see appendix 6.D). In stead of estimating the reservation price for each new good separately, they proceed by estimating a crude demand system for a group of products and then they deduce the effects of new goods from this system. The particular assumptions that are required with this method are constant elasticity of substitution (CES) and absence of changes in the quality of products. Variants of the Feenstra approach are provided by Balk (2000).

The indices presented in this chapter are not bilateral but, like Drobisch' index, unilateral, meaning that they are computed for single sets of goods only. The difference in, for example, prices between two sets of goods can be determined simply by comparing the price index of one set with the price index of the other set. The important advantage of unilateral indices is that the types of products do not have to be the same in both sets. Changes in product variety therefore do not lead to biases in unilateral indices.

Besides being unilateral, the indices presented here are also 'axiomatic'. The distinction between axiomatic and non-axiomatic (or 'economic') indices is that the latter are only meaningful if the assumptions on which they are based, such as optimizing behavior, are valid.<sup>3</sup> The indices are shown to be consistent with two equivalent sets of axioms.

The variety index presented in chapter 5 plays a central role in this chapter as well. Not only is the variety index a component of the quantity index proposed below, it is also an ingredient in the decomposition of productivity and the cost of living.

The chapter proceeds as follows. In section 6.2, two sets of axioms and a new set of indices are proposed. Section 6.3 presents the results of Monte-Carlo simulations comparing the proposed indices with Laspeyres, Paassche and Fisher indices. Section 6.4 contains applications of the proposed indices to the measurement of productivity and the 'cost of living' albeit that these applications are limited to homothetic functions. Additional results are presented for CES functions. Concluding remarks can be found in section 6.5.

## 6.2 Variety robust indices

A common approach to indices is that an index should measure a characteristic of a set of products as if the set would only contain one type of product, that is, as if the set of products is homogenous. For example, a quantity index measured over a heterogenous set of products should be comparable to the total quantity for a homogenous set of products. According to this approach, a set of complementary indices must be consistent with the following axiom.

**Axiom 1** *For a set  $G$  containing  $g$  different types of products with prices  $p_i \in G$  and quantities  $x_i \in G$  a set of a price index  $P$  and a quantity index  $Q$  must satisfy the following conditions:*

1.  $P$  is monotonically increasing in  $p_i \forall i \in G$  given  $\{x_j \in G\}, \{p_j \in G/j\}$

2.  $Q$  is monotonically increasing in  $x_i \forall i \in G$  given  $\{p_{j \in G}\}, \{x_{j \in G/i}\}$
3.  $E \equiv \sum_i p_i x_i = PQ$  (product test)
4.  $P$  is homogenous of degree 1 in  $\{p_{i \in G}\}$  and of degree 0 in  $\{x_{i \in G}\}$
5.  $Q$  is homogenous of degree 1 in  $\{x_{i \in G}\}$  and of degree 0 in  $\{p_{i \in G}\}$
6.  $P = p_1$  if  $g = 1$
7.  $Q = x_1$  if  $g = 1$

The homogeneity conditions are necessary for correct handling of changes in the units of measurement. For example, when prices are measured in dollars, then the price index should equal the price index based on euro prices times the (real) exchange rate. The quantity index should not be affected by the chosen currency. The last two conditions require consistency when the set is homogenous.

In stead of requiring that indices should measure characteristics of a set of products as if the set is homogenous, one can also require that indices measure characteristics of a set as if all product types have identical prices and quantities, that is, as if the product types are symmetric. This approach is reflected in the following variant of the previous axiom.

**Axiom 2** *A set of a price index  $P$ , a quantity index  $X$ , and an index of product variety  $N$ , must satisfy the following conditions:*

1.  $P$  is monotonically increasing in  $p_i \forall i \in G$  given  $\{x_{j \in G}\}, \{p_{j \in G/i}\}$
2.  $X$  is monotonically increasing in  $x_i \forall i \in G$  given  $\{p_{j \in G}\}, \{x_{j \in G/i}\}$
3.  $E = NPX$  (product test)
4.  $P$  is homogenous of degree 1 in  $\{p_{i \in G}\}$  and of degree 0 in  $\{x_{i \in G}\}$
5.  $X$  is homogenous of degree 1 in  $\{x_{i \in G}\}$  and of degree 0 in  $\{p_{i \in G}\}$
6.  $N$  is homogenous of degree 0 in  $\{x_{i \in G}\}$  and of degree 0 in  $\{p_{i \in G}\}$
7.  $P = \bar{p}$  if  $\forall i : p_i = \bar{p}$
8.  $X = \bar{x}$  if  $\forall i : x_i = \bar{x}$
9.  $N = g$  if  $\forall i : (p_i = \bar{p} \text{ and } x_i = \bar{x})$

The last three conditions require consistency when the product types are symmetric. In particular, the index  $N$  is an index of product variety that reduces to the number of product types if all types have identical prices and quantities. The index  $N$  is the antilog of Shannon's (1948) entropy that has been discussed in chapter 5.

**Theorem 4** *The indices*

$$N \equiv \prod_G w_i^{-w_i} \quad (6.1)$$

$$P \equiv \prod_G p_i^{w_i} \quad (6.2)$$

$$X \equiv \prod_G x_i^{w_i} \quad (6.3)$$

$$w_i \equiv \frac{p_i x_i}{\sum_G p_j x_j} \quad (6.4)$$

are consistent with axiom 2.

**Proof.** <sup>4</sup>

Conditions 1, 2 Follow directly from the definitions of  $P$  and  $X$ .

$$\begin{aligned} \text{Condition 3} \quad N P X &= \prod_i \left( \frac{1}{w_i} p_i x_i \right)^{w_i} \\ &= \prod_i \left( \frac{p_i x_i}{p_i x_i / \sum_j p_j x_j} \right)^{w_i} \\ &= \prod_i E^{w_i} = E \end{aligned}$$

Condition 4 The  $w_i$  are homogenous of degree 0 in prices and quantities:  $\frac{\lambda p_i x_i}{\sum \lambda p_j x_j} = \frac{p_i x_i}{\sum p_j x_j} = w_i$   
 $\Rightarrow P$  is homogenous of degree 0 in quantities.  
 The degree of homogeneity of  $P$  in prices is 1:  
 $\prod (\lambda p_i)^{w_i} = \prod \lambda^{w_i} \prod p_i^{w_i} = \lambda P$

Conditions 5, 6 Analogous to condition 4.

$$\text{Condition 7} \quad P = \prod \bar{p}^{w_i} = \bar{p} \prod 1^{w_i} = \bar{p}$$

Condition 8 Analogous to condition 7.

$$\begin{aligned} \text{Condition 9} \quad &\text{if } \forall i : (p_i = \bar{p} \text{ and } x_i = \bar{x}) \text{ then } w_i = \frac{1}{n} \\ &\Rightarrow N = \left( \left( \frac{1}{n} \right)^{-\frac{1}{n}} \right)^n = n \end{aligned}$$

■

Combination of the third conditions of axiom 1,  $E = PQ$ , and axiom 2,  $E = PNX$ , implies that  $Q = NX$  should hold for the two axioms to be consistent with each other.

**Theorem 5**  $P$  and  $Q \equiv NX$  are consistent with axiom 1.

**Proof.** All conditions on  $P$  are identical to those under axiom 2 and therefore they all hold. Condition 2 is satisfied as

$$Q = \prod_G \left( \frac{x_i}{w_i} \right)^{w_i} = \prod_G \left( \frac{\sum_{j \in G} p_j x_j}{p_i} \right)^{w_i} \quad (6.5)$$

and both  $\sum_{j \in G} p_j x_j$  and  $w_i$  are increasing in  $x_i$ . Condition 3 holds as  $E = P(NX) = PQ$ . Compliance with conditions 5 and 7 follows from the properties of  $N$  and  $X$ . ■

Diewert (2001) gives a survey of properties that are considered to be desirable for price and quantity indices. The proposed indices possess most of these properties, but not all of them. Some of the desirable properties mentioned in this survey apply to the weights and can only be possessed by bilateral indices.



### 6.3 Simulation results

In order to get an indication of how the indices  $P$  and  $Q$  relate to established indices, some Monte-Carlo simulations have been performed. The simulations are based upon the indexDixit-StiglitzDixit-Stiglitz framework of monopolistic competition.<sup>5</sup> The preferences of consumers are implied by a CES consumption function  $C = (\sum_i y_i^\gamma)^{\frac{1}{\gamma}}$ . Each good is manufactured by only one firm with the use of labor,  $y_i = b_i l_i$ , where the 'technical efficiency' of firms is heterogenous. In this setting of monopolistic competition, prices are markups over the wage rate,  $p_i = w/(\gamma b_i)$ . The consumption index,  $C$ , is the numeraire.

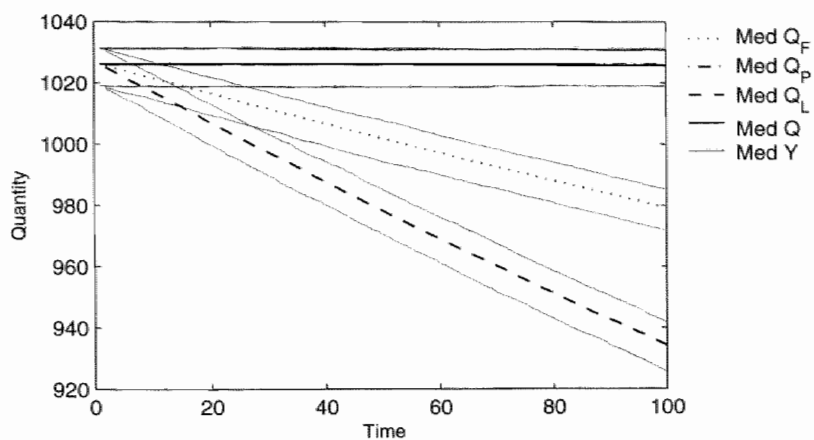
Two sets of simulations have been performed. In the first set the number of firms increases by one each period; in the second set the number of firms decreases by one each period. The  $b_i$  are distributed lognormally across firms with mean 1 and variance 0.01 but are constant over time. The number of firms,  $n$ , starts from 1000, the total labor force  $L = \sum_i l_i$  is 1000 (constant over time), and  $\gamma$  is set to 0.75. Each of the two simulation setups is replicated 200 times. The results in levels for quantity and price are displayed in figures 6.1 to 6.4. The thick lines correspond to the median of the index, the thin lines correspond to the 5% and 95% quantiles implying that 90% of the replications lie between the thin lines.

Besides  $Q$ , the following other quantity indices are reported in figure 6.1: Laspeyres ( $Q_L$ ), Paassche ( $Q_P$ ), Fisher ( $Q_F$ ) and the ordinary sum of quantities,  $Y = \sum_i y_i$ .<sup>6</sup> The Paassche quantity index uses current prices as weights and is therefore not sensitive to increases in the set of products as no arbitrary decisions have to be made about the prices of new products in the previous period.

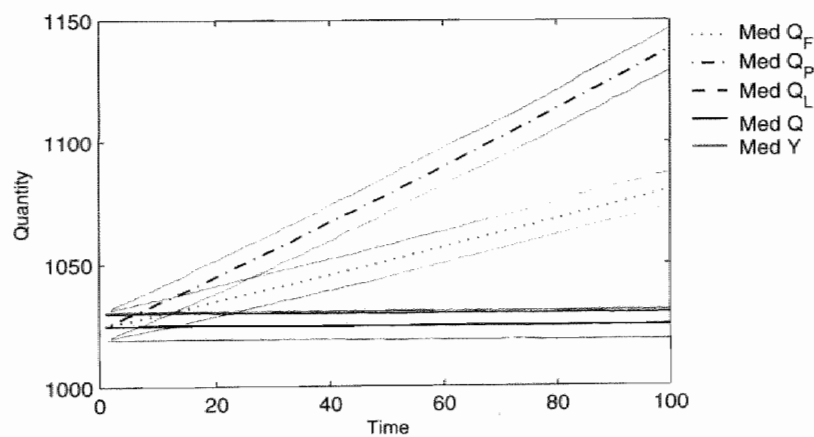
In contrast, the Laspeyres quantity index does rely on prices from the previous period and is therefore biased if the number of goods is increasing. The Fisher index being a geometric average of an unbiased and a biased index is also biased but less so than the Laspeyres index.  $Q$  and  $Q_P$ , the two unbiased indices in this case, remain very close to each other, while  $Q_L$  and  $Q_F$  diverge.

Figure 6.2 applies to the case where the number of product types is decreasing. For contracting product sets, the results are the reverse of those for expanding product sets. With a decreasing number of types, the Paassche index is biased whereas the Laspeyres index is unbiased as the latter relies on previous prices rather than current prices. Now the pair of unbiased indices is  $Q$  and  $Q_L$ . The conclusion to be drawn from figures 6.1 and 6.2 is that  $Q$  is the only index that is unbiased for both expanding and contracting product sets and that the Fisher index is biased in both cases.

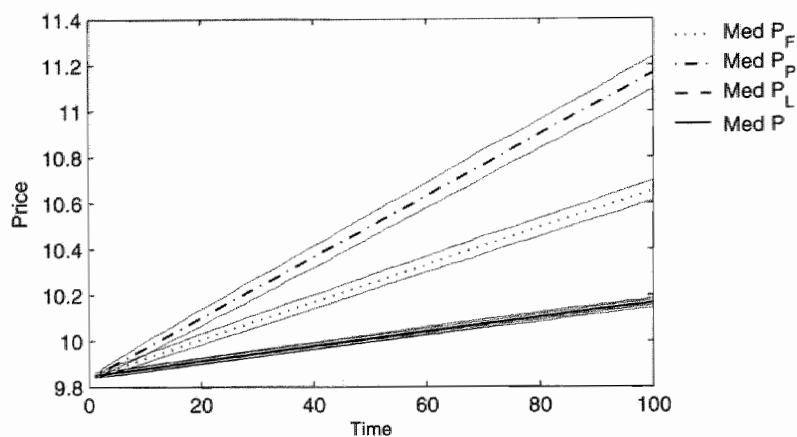
Figures 6.3 and 6.4 show the simulation results for  $P$  and the Paassche, Laspeyres, and Fisher price indices. A first observation is that the number of firms positively affects all price indices. The reason is that the number of product types affects the standard of living despite a constant production capacity. Second, when the number of types rises over time  $P$  grows at approximately the same rate as the Laspeyres index, while  $P$  is close to the Paassche index if the number of types declines over time.  $P$  is the only price index that is unbiased for both expanding and contracting product sets.



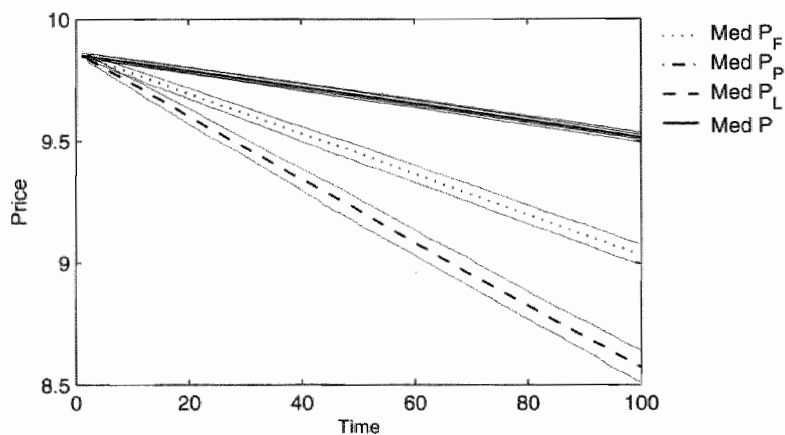
**Figure 6.1.** Simulation results for quantity indices with an increasing number of product types (chain method)



**Figure 6.2.** Simulation results for quantity indices with a decreasing number of product types (chain method)



**Figure 6.3.** Simulation results for price indices with an increasing number of product types (chain method)



**Figure 6.4.** Simulation results for price indices with a decreasing number of product types (chain method)

## 6.4 Applications to economic theory

### 6.4.1 Productivity measurement

From an axiomatic point of view, a measure of total factor productivity growth (TFPG) that is based on the indices of the previous section is easily constructed: simply compare the ratio of the quantity index of the inputs,  $Q_x$ , to that of the outputs,  $Q_y$ , in one year or country to that of another one.

$$TFPG = \ln(Q_y^1/Q_x^1) - \ln(Q_y^0/Q_x^0) \quad (6.6)$$

It is more complicated to show that this TFPG-index is consistent with economic theory. In fact, I will limit my discussion to homothetic production functions with many inputs and just one output. Consider a firm with a continuously differentiable homothetic production function of the form

$$y = h(f(\mathbf{x})), \quad (6.7)$$

where  $y$  is the quantity of output produced,  $\mathbf{x}$  is a vector of input quantities,  $f$  is a homogenous function of degree one, and  $h$  is a continuous, monotonically increasing function of  $f$ .<sup>7</sup> The firm takes the prices of inputs as given.

Profit maximization yields first order conditions that implicitly relate the amounts of inputs demanded with their prices.

$$\max_{x_i} \left\{ p_y(y)y - \sum p_i x_i \right\} \Rightarrow p_i = (1 + \varepsilon_{p_y y}) h' f_i p_y \quad (6.8)$$

Here,  $\varepsilon_{p_y y}$  is the price elasticity of demand.

Cost shares are given by

$$w_i = \frac{p_i x_i}{\sum p_j x_j} = \frac{(1 + \varepsilon_{p_y y}) h' f_i p_y x_i}{\sum (1 + \varepsilon_{p_y y}) h' f_j p_y x_j} = \frac{f_i x_i}{\sum f_j x_j}. \quad (6.9)$$

Euler's theorem allows us to write  $f(\mathbf{x}) = \sum f_i x_i$ . Applying theorem 6 provided in the appendix to the right hand side of this expression,  $f$  can be decomposed into the variety index  $N$  (as defined by equation 6.1), the average quantity index of the inputs  $X$  (as defined by equation 6.3), and an index of the partial derivatives  $F'$ .

$$f(\mathbf{x}) = NF'X \quad (6.10)$$

$$F' \equiv \prod f_i^{w_i}, \quad w_i = \frac{f_i x_i}{\sum f_j x_j} \quad (6.11)$$

The assumption of homotheticity assures that the weights used in the indices equal the cost shares of the corresponding products.

Total factor productivity growth is the log change in  $y/Q$  and can be decomposed as follows.

$$d \ln \left( \frac{y}{Q} \right) = d \ln h(NF'X) - d \ln Q \quad (6.12)$$

$$= \varepsilon_{hf} (d \ln N + d \ln F' + d \ln X) - d \ln Q \quad (6.13)$$

$$= (\varepsilon_{hf} - 1) d \ln Q + \varepsilon_{hf} d \ln F' \quad (6.14)$$

Total factor productivity growth depends on the quantity of inputs used (a scale effect) and on the index of partial derivatives of  $f(\mathbf{x})$  (a technology effect). These

two effects are moderated by  $\varepsilon_{hf}$ , the 'elasticity of scale'. Of course, this decomposition only is exact for marginal changes in  $y$  and  $\mathbf{x}$ .

Usually we are not interested in marginal but in substantial differences in productivity. Allowing for substantial differences the decomposition of TFPG is still exact if  $h$  is homogeneous rather than homothetic. For homogenous functions of degree  $r$ , Euler's theorem states that  $h = \frac{1}{r} \sum h_i x_i$ . Decomposing  $h$  yields  $\frac{1}{r} N H' X$  with the weights  $h_i x_i / \sum h_j x_j$  equaling expenditure shares. TFPG allowing for substantial differences now can be stated as

$$\Delta \ln \left( \frac{y}{Q} \right) = \Delta \ln H', \quad H' \equiv \prod h_i^{w_i}. \quad (6.15)$$

### 6.4.2 CES functions

If we are willing to impose even more structure on the production function and require it to have constant elasticity of substitution between inputs (CES), TFPG can be decomposed even further. The production function given by

$$y = h(\mathbf{x}) = \left( \sum \theta_i x_i^\gamma \right)^{\frac{r}{\gamma}}, \quad (6.16)$$

allows for  $y$  to be expressed in terms of  $N$  and  $X$  using theorem 6.

$$y^{\frac{\gamma}{r}} = \sum \theta_i x_i^\gamma = N \Theta X^\gamma \Rightarrow y = N^{\frac{r}{\gamma}} \Theta^{\frac{r}{\gamma}} X^r \quad (6.17)$$

$$w_i = \frac{\theta_i x_i^\gamma}{\sum \theta_j x_j^\gamma}, \quad \Theta \equiv \prod \theta_i^{w_i} \quad (6.18)$$

Here,  $\Theta$ , a 'quality-index', is the weighted geometric mean of the  $\theta_i$ . It remains to be demonstrated that these weights equal the cost shares. Profit maximization implies that the price of good  $i$  equals  $(1 + \varepsilon_{pyy}) r y^{1-\frac{r}{\gamma}} \theta_i x_i^{\gamma-1} p_y$ . Applying this to the cost shares confirms that the weights are indeed correct.

$$w_i = \frac{p_i x_i}{\sum p_j x_j} = \frac{(1 + \varepsilon_{pyy}) r y^{1-\frac{r}{\gamma}} \theta_i x_i^{\gamma} p_y}{\sum (1 + \varepsilon_{pyy}) r y^{1-\frac{r}{\gamma}} \theta_j x_j^{\gamma} p_y} = \frac{\theta_i x_i^\gamma}{\sum \theta_j x_j^\gamma} \quad (6.19)$$

Using  $Q = NX$ , an expression for TFPG can readily be obtained:

$$\Delta \ln \left( \frac{y}{Q} \right) = \Delta \ln \left( N^{\frac{(1-\gamma)r}{\gamma}} \Theta^{\frac{r}{\gamma}} Q^{r-1} \right) \quad (6.20)$$

$$= (r-1) \Delta \ln Q + \frac{(1-\gamma)r}{\gamma} \Delta \ln N + \frac{r}{\gamma} \Delta \ln \Theta \quad (6.21)$$

The first term is the scale effect, the second is the effect due to a change in product variety, and the last term is the effect attributable to changes in quality.

### 6.4.3 Cost of living indices

Cost of living indices (COLI) differ from normal price indices in that they refer to the effect of a change in prices rather than to the change in prices itself. In principle, the cost of living can change for two reasons. First, the prices of consumer goods may change. Second, consumers may appreciate their basket of goods differently than before. This second component requires a COLI to be rooted into

economic theory. Below I will show how the indices proposed in this chapter can be used for the construction of a COLI based on the homothetic utility function

$$u = h(f(\mathbf{x})), \quad (6.22)$$

where  $u$  is the utility of a representative household,  $\mathbf{x}$  is a vector of quantities of consumer goods,  $f$  is a homogenous function of degree one, and  $h$  is a continuous, monotonically increasing function of  $f$ .

The consumer maximizes utility subject to the budget  $E \equiv \sum p_i x_i$ .

$$\max_{x_i} \left\{ h(f(\mathbf{x})) + \lambda \left( E - \sum p_i x_i \right) \right\} \quad (6.23)$$

From the first order condition follows that prices are given by  $p_i = \lambda^{-1} h' f_i$ , with  $\lambda$  being the Lagrange multiplier. Consequently, budget shares can be expressed as in 6.9.

$$w_i = \frac{p_i x_i}{\sum p_j x_j} = \frac{\lambda^{-1} h' f_i x_i}{\sum_j \lambda^{-1} h' f_j x_j} = \frac{f_i x_i}{\sum_j f_j x_j} \quad (6.24)$$

Applying Euler's theorem and the theorem from the appendix, the utility function can be rewritten as  $u = h(F'Q)$ . For utility functions that are homogenous of degree  $r$  this becomes  $u = \frac{1}{r} H'Q$  (see subsection 6.4.1).

A COLI is defined as the unit cost of utility and is decomposed as

$$c \equiv E/u = \frac{PQ}{h(F'Q)} \quad (6.25)$$

$$d \ln c = d \ln P - (\varepsilon_{hf} - 1) d \ln Q - \varepsilon_{hf} d \ln F' \quad (6.26)$$

for the homothetic case, and as

$$c = E/u = \frac{PQ}{\frac{1}{r} H'Q} = rP/H' \quad (6.27)$$

$$\Delta \ln c = \Delta \ln P - \Delta \ln H' \quad (6.28)$$

for the homogenous case. Changes in the cost of living are caused by a change in the price index of consumption goods, by a change in the 'efficiency' of consumption, and by a change in the level of consumption (an 'income-effect').

The equivalence of these results with COLI's based on CES-preferences is easily demonstrated. As has been shown in subsection 6.4.1 the CES function  $u = (\sum \theta_i x_i^\gamma)^{\frac{1}{\gamma}}$  can be written as  $N^{\frac{r}{\gamma}} \Theta^{\frac{r}{\gamma}} X^r$ . The COLI belonging to this function is easy to derive.

$$c = E/u = \frac{NPX}{N^{\frac{r}{\gamma}} \Theta^{\frac{r}{\gamma}} X^r} = N^{\frac{(\gamma-1)r}{\gamma}} \Theta^{\frac{-r}{\gamma}} PQ^{1-r} \quad (6.29)$$

$$\Delta \ln c = \frac{(\gamma-1)r}{\gamma} \Delta \ln N - \frac{r}{\gamma} \Delta \ln \Theta + (1-r) \Delta \ln Q + \Delta \ln P \quad (6.30)$$

The cost of living declines with increases in product variety and product quality since  $0 < \gamma < 1$ , but it rises with increases the price index of consumption goods and with the quantity consumed (if  $r < 1$ ).

## 6.5 Conclusion

The indices proposed in this chapter can be used for the comparison of sets of goods that do not contain the same types of products. Changes in product variety do not lead to biases in the indices. Moreover, the indices do not rely on assumptions and they unconditionally pass the product test,  $E = PQ$ .

It has been shown that the indices can be used for the decomposition of productivity growth into scale and non-scale effects provided that there is a single output and that the production function is homothetic. In a similar fashion, the indices allow for the decomposition of changes in the cost of living into the effect of changes in the price index of consumer goods and the effect of changes in the 'efficiency' in consumption, conditional upon the utility function being homothetic.

In the special case of CES functions, the indices allow for the exact separation of the effects of product variety and product quality on productivity and the cost of living. For this reason, and because the CES function need not be homogenous of degree one, the proposed indices can be considered an improvement over the indices proposed by Feenstra.

The variety index of chapter 5 complements the price and average quantity indices presented above. Decompositions of productivity and the cost of living reveal that product variety as measured by this index has effects that are distinct from the effects attributable to changes in the aggregate quantity index.

## 6.A Multiplicative decomposition of additive indices

**Theorem 6** *Any index  $u$  measured over a set  $N$  that satisfies*

$$u_N = \sum_{i \in N} u_i \quad (6.31)$$

$$u_i = \prod_{j \in M} x_{ij}^{a_j} \quad (6.32)$$

*can be decomposed into indices  $X_j$  for all  $j \in M$  such that*

$$u_N = \exp[H] \prod_{j \in M} X_j^{a_j} \quad (6.33)$$

*where*

$$H = - \sum_{i \in N} w_i \ln w_i \quad (6.34)$$

$$X_j = \prod_{i \in N} x_{ij}^{w_i} \quad (6.35)$$

$$w_i \equiv \frac{u_i}{u_N}. \quad (6.36)$$

**Proof.**

$$\ln(u_N) = H + \sum_{j \in M} \ln(X_j^{a_j}) \quad (6.37)$$

$$= \sum_{i \in N} w_i \ln \frac{1}{w_i} + \sum_{i \in N} w_i \ln \left[ \prod_{j \in M} x_{ij}^{a_j} \right] \quad (6.38)$$

$$= \sum_{i \in N} w_i \ln \left[ \frac{1}{w_i} \prod_{j \in M} x_{ij}^{a_j} \right] \quad (6.39)$$

$$= \sum_{i \in N} w_i \ln \left[ \frac{\prod_{j \in M} x_{ij}^{a_j}}{\prod_{j \in M} x_{ij}^{a_j} / \sum_{i \in N} \prod_{j \in M} x_{ij}^{a_j}} \right] \quad (6.40)$$

$$= \ln \left[ \sum_{i \in N} \prod_{j \in M} x_{ij}^{a_j} \right] \sum_{i \in N} w_i = \ln \left[ \sum_{i \in N} u_i \right] \quad (6.41)$$

■

## 6.B Application to CRESH functions

The homothetic constant ratios of elasticity of substitution (CRESH) production function devised by Hanoch (1971) differs from the general homothetic functions of section 6.4 in that its output is an implicit function of the inputs. Below it is shown that for the CRESH function results can be found similar to those of section 6.4.

The CRESH function is defined as

$$G(y, \mathbf{x}) = \sum_{i=1}^n \theta_i [x_i g(y)]^{\delta_i} - 1 \equiv 0. \quad (6.42)$$

The derivative of  $y$  with respect to  $x_j$  can be found using the implicit function theorem.

$$\frac{\partial y}{\partial x_j} = - \frac{G_{x_j}}{G_y} = \frac{\theta_j \delta_j x_j^{\delta_j - 1} g(y)^{\delta_j}}{\sum_{i=1}^n \theta_i \delta_i x_i^{\delta_i} g(y)^{\delta_i - 1} g'} \quad (6.43)$$

The first order condition for profit maximization ensures that  $p_j = (1 + \varepsilon_{p_y y}) \frac{\partial y}{\partial x_i} p_y$ . This leads to the following expression for expenditure shares:

$$w_i = \frac{\frac{\theta_i \delta_i x_i^{\delta_i - 1} g(y)^{\delta_i}}{\sum_i \theta_j \delta_j x_j^{\delta_j} g(y)^{\delta_j - 1} g'} x_i}{\sum_j \frac{\theta_j \delta_j x_j^{\delta_j - 1} g(y)^{\delta_j}}{\sum_k \theta_k \delta_k x_k^{\delta_k} g(y)^{\delta_k - 1} g'} x_j} \quad (6.44)$$

$$= \frac{\theta_i \delta_i x_i^{\delta_i} g(y)^{\delta_i}}{\sum_j \theta_j \delta_j x_j^{\delta_j} g(y)^{\delta_j}}. \quad (6.45)$$

We use these expenditure shares for the construction of an index of the derivatives  $\partial y / \partial x_j$ .

$$F' \equiv \prod_j \left( \frac{g(y)}{g' x_j} \frac{\theta_j \delta_j x_j^{\delta_j} g(y)^{\delta_j}}{\sum_{i=1}^n \theta_i \delta_i x_i^{\delta_i} g(y)^{\delta_i}} \right)^{w_j} \quad (6.46)$$

$$= \frac{g(y)}{g' X \sum_i \theta_i \delta_i x_i^{\delta_i} g(y)^{\delta_i}} \prod_j \left( \theta_j \delta_j x_j^{\delta_j} g(y)^{\delta_j} \right)^{w_j} \quad (6.47)$$



The summation below the line can be decomposed using the same weights as above.

$$\sum_i \theta_i \delta_i x_i^{\delta_i} g(y)^{\delta_i} = N \prod_i \left( \theta_i \delta_i x_i^{\delta_i} g(y)^{\delta_i} \right)^{w_i} \quad (6.48)$$

Insert this expression back into  $F'$  to get an index for productivity that is of identical form as that in subsection 6.4.1.

$$F' = \frac{g(y)}{g' X N \prod_i \left( \theta_i \delta_i x_i^{\delta_i} g(y)^{\delta_i} \right)^{w_i}} \prod_j \left( \theta_j \delta_j x_j^{\delta_j} g(y)^{\delta_j} \right)^{w_j} \quad (6.49)$$

$$\frac{y}{Q} = \varepsilon_{gy}^{-1} F' \quad (6.50)$$

The COLI corresponding to the CRESH function is

$$c = E/y = \frac{NPX}{\varepsilon_{gy}^{-1} F' Q} = \varepsilon_{gy} P/F'. \quad (6.51)$$

## 6.C Comment on impossibility theorems

Two impossibility results have been established that suggest that a set of unilateral indices cannot simultaneously satisfy a basic set of conditions. The first impossibility result is due to Eichhorn (1978, p. 144) and applies when  $P = P(\mathbf{p})$  and  $Q = Q(\mathbf{x})$ ,  $\mathbf{p}$  and  $\mathbf{x}$  being vectors of  $g$  prices and quantities, respectively. The proof of the result stated below, is a modified version of that of Eichhorn. Suppose that  $\mathbf{p} = \mathbf{x} = \mathbf{1}$ , then Eichhorn required  $P(\mathbf{1}) = a$  and  $Q(\mathbf{1}) = b$  where  $a$  and  $b$  are constants. If  $P$  and  $Q$  are to satisfy  $PQ = \mathbf{p}'\mathbf{x}$  then it must hold that  $ab = g$  when  $\mathbf{p} = \mathbf{x} = \mathbf{1}$ . This condition cannot hold in general. Eichhorn makes a mistake in requiring  $Q(\mathbf{1}) = b$  as  $Q$  is a measure of aggregate quantity and therefore does not only rely on  $x$  but also on  $g$ , i.e.  $Q(\mathbf{1}) = b(g)$ . For  $\mathbf{p} = \mathbf{x} = \mathbf{1}$  the indices proposed in this chapter reduce to  $P = 1$  and  $Q = g$  such that  $PQ = \mathbf{p}'\mathbf{x} = g$ . The impossibility result is not robust to a small and justified change in the conditions that the indices should meet.

The second impossibility result is also due to Eichhorn (1978, p. 144-145) and has been refined by Diewert (1993, p. 9). Suppose that  $P = P(\mathbf{p}, \mathbf{x})$  and  $Q = Q(\mathbf{p}, \mathbf{x})$  and also that the homogeneity conditions of axiom 1 have been met. Additionally, suppose that  $P$  and  $Q$  are invariant to the units of measurement of the quantities such that

$$P(p_1 d_1, \dots, p_g d_g; x_1/d_1, \dots, x_g/d_g) \equiv P(\mathbf{p}, \mathbf{x}, \mathbf{d}) = P(\mathbf{p}, \mathbf{x}) \quad (6.52)$$

where  $d_i$  is a product specific change in the unit of measurement. If we take  $d_i = x_i$  then

$$P(\mathbf{p}, \mathbf{x}) = P(p_1 x_1, \dots, p_g x_g; 1, \dots, 1) \quad (6.53)$$

which is inconsistent with the requirement that  $P$  is homogenous of degree 0 in  $\mathbf{x}$ .

It can easily be checked that the proposed indices are not invariant to the units of measurement (unless  $\forall i : d_i = d$ ). This obviously invalidates the proposed quantity index as a direct measure of consumer-welfare. However, this property is desirable if the index is to be used as a pure quantity index: ordinary prices and quantities also are not invariant to their units of measurements. A pound of sugar is not traded at the same price as a kilogram of sugar. A consequence of this property is that when relative changes in quantities vary across products, units of measurement also matter for the relative change in the quantity index.

## 6.D Feenstra's variety index

Feenstra's variety index applies only to CES functions that homogenous of degree one. Suppose  $r = 1$  and  $p_y y = \sum p_i x_i$ , then  $p_y$  can be written as  $p_y(\theta, \mathbf{p})$  using equation 6.19.

$$\frac{p_i}{p_y} = y x_i^{-1} \frac{\theta_i x_i^\gamma}{\sum \theta_j x_j^\gamma} \quad (6.54)$$

$$p_i^{\frac{\gamma}{\gamma-1}} p_y^{\frac{1}{1-\gamma}} \theta_i^{\frac{1}{1-\gamma}} = \left( \sum \theta_i x_i^\gamma \right)^{\frac{1}{\gamma}} p_i x_i \quad (6.55)$$

$$\sum p_i^{\frac{\gamma}{\gamma-1}} p_y^{\frac{1}{1-\gamma}} \theta_i^{\frac{1}{1-\gamma}} = \left( \sum \theta_i x_i^\gamma \right)^{\frac{1}{\gamma}} E = \frac{E}{y} = p_y \quad (6.56)$$

$$p_y = \left( \sum \theta_i^{\frac{1}{1-\gamma}} p_i^{\frac{\gamma}{\gamma-1}} \right)^{\frac{\gamma-1}{\gamma}} \quad (6.57)$$

Rewrite the weights in terms of prices for the set  $G$  of available product types.

$$w_{it}(G_t) = \frac{p_{it} x_{it}}{p_{yt} y_t} = p_{it}^{\frac{\gamma}{\gamma-1}} p_{yt}^{\frac{1}{1-\gamma}} \theta_{it}^{\frac{1}{1-\gamma}} \quad (6.58)$$

The set of products available in periods  $t$  and  $t-1$  is  $G_t^* \equiv G_t \cap G_{t-1}$ . Let some  $H_t \subseteq G_t^*$ .

$$w_{it}(G_t) = w_{it}(H_t) \lambda(G_t, H_t) \quad (6.59)$$

$$\lambda(G_t, H_t) \equiv \frac{\sum_{H_t} p_{it} x_{it}}{\sum_{G_t} p_{it} x_{it}} \quad (6.60)$$

Combine the latter expression with 6.58 and solve for  $p_y$  to get

$$\frac{p_{yt}}{p_{yt-1}} = \frac{w_{it}(G_t)^{\frac{1-\gamma}{\gamma}} p_{it} \theta_{it}^{\frac{-1}{\gamma}}}{w_{it-1}(G_{t-1})^{\frac{1-\gamma}{\gamma}} p_{it-1} \theta_{it-1}^{\frac{-1}{\gamma}}} \quad (6.61)$$

$$= \left( \frac{w_{it}(H_t) \lambda(G_t, H_t)}{w_{it-1}(H_t) \lambda(G_{t-1}, H_t)} \right)^{\frac{1-\gamma}{\gamma}} \left( \frac{\theta_{it}}{\theta_{it-1}} \right)^{\frac{-1}{\gamma}} \frac{p_{it}}{p_{it-1}}. \quad (6.62)$$

Take the geometric average using Sato-Vartia weights over set  $H$ ,  $s_{it}(H_t)$ .

$$\frac{p_{yt}}{p_{yt-1}} = P_{SV}(H_t) \Theta_{SV}(H_t)^{\frac{-1}{\gamma}} \left( \frac{\lambda(G_t, H_t)}{\lambda(G_{t-1}, H_t)} \right)^{\frac{1-\gamma}{\gamma}} \quad (6.63)$$

$$P_{SV}(H_t) \equiv \prod_{H_t} \left( \frac{p_{it}}{p_{it-1}} \right)^{-s_{it}(H_t)}, \quad \Theta_{SV}(H_t) \equiv \prod_{H_t} \left( \frac{\theta_{it}}{\theta_{it-1}} \right)^{-s_{it}(H_t)} \quad (6.64)$$

$$s_{it}(H_t) \equiv \frac{\frac{w_{it}(H_t) - w_{it-1}(H_t)}{\ln w_{it}(H_t) - \ln w_{it-1}(H_t)}}{\sum_{H_t} \frac{w_{it}(H_t) - w_{it-1}(H_t)}{\ln w_{it}(H_t) - \ln w_{it-1}(H_t)}} \quad (6.65)$$

The price ratio depends on the Sato-Vartia price index, a 'Sato-Vartia quality index', and the ratio of  $\lambda$ 's.<sup>8</sup>

Define total factor productivity growth, TFPG, as the relative change in output minus the relative change in expenditure deflated by the Sato-Vartia price

index for the subset  $H$ .

$$\text{TFPG} \equiv \Delta \ln y_t - \ln \left( \frac{E_t/E_{t-1}}{P_{SV}(H_t)} \right) = -\ln \left( \frac{p_{yt}/p_{yt-1}}{P_{SV}(H_t)} \right) \quad (6.66)$$

$$= \frac{1}{\gamma} \ln \Theta_{SV}(H_t) - \frac{1-\gamma}{\gamma} \ln \left( \frac{\lambda(G_t, H_t)}{\lambda(G_{t-1}, H_t)} \right) \quad (6.67)$$

The ratio  $\lambda(G_t, H_t)/\lambda(G_{t-1}, H_t)$  is known as Feenstra's variety index. Because expenditure on the set  $G$  is deflated by a price index defined over  $H \subseteq G$ , Feenstra's measure TFPG is just an approximation of total factor productivity growth. The Feenstra index is therefore not even exact in the case of a CES function that is homogenous of degree one.

## Notes

1. Erwin Diewert brought to my attention the existence of two impossibility theorems regarding unilateral indices. A comment on these theorems can be found in appendix 6.C.
2. Hausman (2003) discerns three types of bias in bilateral indices: bias caused by changes in expenditure shares induced by changes in prices (substitution bias), bias caused by the introduction of new goods, bias stemming from changes in the quality of goods, and bias due to changes in shopping patterns (outlet bias). Note that the bias caused by new products is an extreme form of substitution bias. The indices presented in this chapter do not suffer from the first two types of bias; the COLI and productivity indices of section 6.4 take into account quality changes (for CES consumption and production functions).
3. Strict application of this definition implies that the Laspeyres and Paasche indices are non-axiomatic as they rely on the assumption that the product types in the reference set are the same as those in the comparison set.
4. A generalized version of the proof corresponding to condition 3 (product test) can be found in appendix 6.B.
5. The number of firms is exogenous and there are no fixed costs of production.
6. All bilateral indices are computed according to the chain method. Using a fixed base would lead to a larger bias.
7. It is shown in the appendix that the results in this section also apply to the CRESH function (Hanoch 1971), which is implicitly homothetic.
8. Feenstra does not consider quality changes, such that originally  $\Theta_{SV} = 1$ .

## Two models of economic growth

### 7.1 Introduction

The literature on endogenous growth has relied on experimenting with specifications for intertemporal knowledge spillovers in order to explain empirical trends like the rising share of researchers in the workforce. A consequence of this rather narrow approach has been that the debates between contributors to this literature have concentrated on parameter settings and functional forms related to knowledge spillovers (see chapter 4 for a review). The two models presented in this chapter are a response to this increasingly metaphysical debate. In the first model, knowledge spillovers are omitted altogether, shifting the focus to education. The model demonstrates how growth in the average level of education can substitute for population growth as a determinant of growth in product variety. This theoretical prediction will be tested in chapter 10. Additionally, it shows how changes in the distribution of educational attainment affect the economy.

The second model contains an alternative kind of intertemporal knowledge spillovers. In particular, knowledge spillovers cause each newly invented product to be an imperfect improvement over the existing products. The model abandons the customary dichotomy that separates quality improvements from variety expansion. In stead, it covers the spectrum of which quality improvements and variety expansion are limiting cases.

Section 7.2 contains a model that links the average level of schooling and schooling inequality to economic growth. The second model, in which new goods are imperfect improvements over old goods, is treated in section 7.3. The overall results and implications are summarized in section 7.4.

### 7.2 Schooling inequality and growth

#### 7.2.1 Introduction

One trend discussed in chapter 3 is the rise of tertiary education in the OECD (figure 3.5). It is common sense to assume that a worker with a higher level of education is likely to be more productive than a worker with a lower level of education. Indeed, at least for the United States there is overwhelming evidence that higher skills lead to higher wages (see Acemoglu 1998, 2002 and Goldin and Katz

2000, 2001). At the aggregate level, a better educated population is therefore expected to translate into a more productive workforce. This implies that growth in the level of education affects economic growth in much the same way as population growth does.

Another empirical trend presented in chapter 3 is the decline in schooling inequality in the OECD. Figure 3.7 shows that the shares of the population with primary, secondary or tertiary education as the highest level of educational attainment have converged, making the distribution of educational attainment more uniform. This trend has also been reported by Ram (1990), who finds that there exists an inverse relationship between schooling inequality and the average years of schooling if the average years of schooling in a country exceeds seven.

The model presented in this section demonstrates the effects of exogenous education growth on consumption growth. Advances in schooling induce economic growth, not only in the trivial way of raising productivity per worker, but also because they stimulate research in the same way as population growth does. Because of the latter effect, consumption per capita grows at a higher rate than the productivity of production workers does. This results has been established earlier by Arnold (1998).

A second feature of the model is that workers differ in their education, or, more precisely, they differ in their skills, but not in the customary low-skill – high-skill dichotomy. Instead, skills are continuously distributed over workers. There are two kinds of jobs in the economy: jobs in the production sector and jobs in the research sector. In both types of jobs, workers with higher skills are more productive and receive a higher income. What distinguishes the two jobs from each other is that productivity is more sensitive to skills in research than it is in production.

If the distribution of skills changes, product variety will be affected in two ways. First, the average productivity of researchers might change. Second, the proportion of researchers in the workforce might be affected. A change in the distribution of skills influences the proportion of researchers in the workforce through two channels. First, a change in the skill level of a worker changes the comparative advantage she has in her current job-type. A worker employed in the production sector, for example, might be more inclined to choose a job in research if her skills increase. I will refer to this as the ‘skill’ effect.

Second, a change in the distribution of skills will alter the amounts of aggregate human capital available for production and for research. In general this will lead to either over- or under-investment in research, such that the wage rates of both sectors will change. The change in wages induces a change in the proportion of researchers. I will refer to this as the ‘wage’ effect.

A mean-preserving change in the distribution of skills will trigger both ‘skill’ and ‘wage’ effects. The direction of the two effects depends on parameter settings. Furthermore, the effects may or may not work in the same directions, such that a change in schooling inequality can have a variety of steady state effects – amongst which a rise in the proportion of researchers.

When parameters are such that a reduction in inequality leads to a rise in the proportion of researchers, a greater variety of products is not guaranteed. Although the reduction in inequality might induce more workers to become a researcher, it also has a negative effect on their productivity – provided that average skills are kept constant. The latter effect follows almost by definition: a mean-preserving decline in skill inequality can only be achieved by ‘transferring’ skills from high skilled workers to low skilled workers.

Several authors have taken the rise in educational attainment as a starting point for explaining economic growth. The most well-known contributions in this direction are the endogenous growth models by Lucas (1988). After Lucas, the theoretical literature of the 1990s – with the exception of Barro, Mankiw, and Sala-i Martin (1995) – has largely ignored the effects of education on economic growth. Interest in the topic revived with the papers by Bils and Klenow (2000) and Jones (2002). In the model by Bils and Klenow the capability of teachers is larger when they are better educated themselves. In this way, young teachers will be more capable than old ones and, consequentially, human capital will accumulate and productivity will grow. The model by Jones builds on that of Bils and Klenow but is augmented with intertemporal knowledge spillovers.

The setup of the model presented below differs from that of Jones (2002) in several respects. First, the model allows for changes in the distribution of skills whereas all workers are equally capable in Jones' model. Second, Jones treats the process of education in some detail, while here educational attainment is exogenous.<sup>1</sup> Third, Jones does not provide a micro-foundation for his model. Finally, Jones takes into account knowledge spillovers.<sup>2</sup>

With regard to the explanation for the rising share of research in employment, the model presented below deviates from Jones' model in that a constant growth rate in average education does not cause the proportion of researchers to grow over time. Instead, the rising share of research is shown to be a possible consequence of a reduction in schooling inequality.

As the current section deals with the consequences of changes in the distribution of skills, it is also related to the literature on skill biased technological change and to the literature on job assignment and occupational choice. The first strand of literature has been surveyed by Acemoglu (2002). Although the model presented here does show that a rise in the 'college-premium' can be a consequence of a more equal distribution of skills, it does not really fit into this literature because it does not incorporate skill biased technological change. Laitner (2000) studies the relation between the distribution of 'natural abilities' and wage inequality using a model with unbiased – but exogenous – technological change.

The assignment of jobs to workers is very simple in the model presented below. There are just two kinds of jobs and skills are one-dimensional. The literature on assignment and occupational choice contains many more advanced configurations (Sattinger 1993). The approach followed by Teulings (1995) is particularly interesting in the context of the model presented here. In his model both the complexity of jobs and the skills of workers have a continuous distribution.

After the basic model has been presented in subsection 7.2.2, its steady state will be solved for in subsection 7.2.3. Subsection 7.2.4 contains a discussion of the effects of growth in average skills and changes in the distribution of skills.

## 7.2.2 The model

Let  $\mathcal{L}$  be the set of all workers and order the workers according to their skills. Skills, denoted by  $k$ , are one-dimensional, implying that they reflect some general notion of intelligence or capability. The workers are indexed from 0 to  $L$ , where  $L$  is both the worker with highest skills and the total amount of labor. The skills of worker  $i$  depend on his relative ranking,  $i/L$ , and on the parameters  $s$  and  $\sigma$  ( $s, \sigma > 0$ ).

$$k(i) = (\sigma + 1) s (i/L)^\sigma \quad (7.1)$$

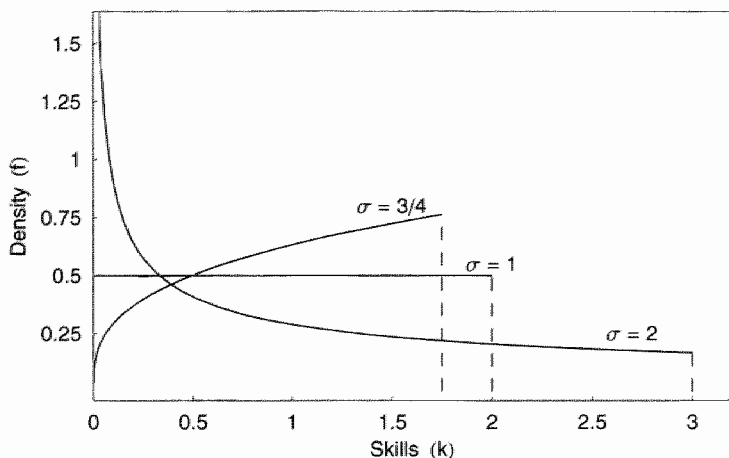


Figure 7.1. Density plots for skill distributions

Integration of  $k(i)$  over  $\mathcal{L}$  shows that  $s$  is simply the average level of skills.

$$\frac{1}{L} \int_0^L (\sigma + 1) s (i/L)^\sigma di = \left[ s \frac{i^{\sigma+1}}{L^{1+\sigma}} \right]_0^L = s \quad (7.2)$$

The chosen specification of  $k$  has the advantage that changes in  $\sigma$  do not affect average skills. The shape of the distribution of  $k$  can be altered by varying  $\sigma$  without affecting the mean of the distribution. This distribution can be derived in a straightforward manner by solving equation 7.1 for  $i/L$ , which yields the cumulative distribution, and differentiating the result with respect to  $k$ .<sup>3</sup>

$$\frac{i}{L} = \frac{1}{[(\sigma + 1) s]^{\frac{1}{\sigma}}} k(i)^{\frac{1}{\sigma}} \quad (7.3)$$

$$f(k) = \frac{1}{\sigma [(\sigma + 1) s]^{\frac{1}{\sigma}}} k^{\frac{1-\sigma}{\sigma}} \quad (7.4)$$

In this last expression,  $f$  is the density function. Figure 7.1 shows the distributions of  $k$  for three values of  $\sigma$ , having  $s = 1$ . The figure illustrates that although the function in 7.1 is quite simple, it still allows for a reasonable flexibility in the distribution of  $k$ .

A larger  $\sigma$  causes the maximum skill level to increase, even though  $s$  is fixed. A rise in  $\sigma$  therefore widens the gap between the workers with minimal skills and the workers with maximal skills. This follows directly from  $k(L) = (\sigma + 1) s$ . Besides affecting the domain of  $k$ ,  $\sigma$  also affects its variance.

$$\text{var}(k) = \int_0^{(\sigma+1)s} k^2 f(k) dk - s^2 = \frac{\sigma s^2}{2 + \sigma^{-1}} \quad (7.5)$$

The variance of  $k$  is clearly increasing in  $\sigma$ . A rise in  $\sigma$  not only widens the skill-gap between workers 0 and  $L$ , it also raises the variance of the skill distribution. These two properties of the skill distribution make  $\sigma$  a reasonably appropriate measure of inequality and whenever I mention 'inequality' below, I implicitly refer to  $\sigma$ .

Consumers have the usual CES preferences over  $n$  symmetric goods.<sup>4</sup> The amount produced of each good is  $x$ , and the production of  $x$  requires an amount of human capital equivalent to  $H_x/n$  (human capital will be defined later on). Aggregate consumption is therefore a function of  $n$  and  $H_x$ .

$$C = n^{\frac{1-\gamma}{\gamma}} H_x \quad (7.6)$$

Here,  $\gamma$  is the same parameter as in equation 4.8. The flow of new goods depends on the amount of human capital available for research,  $H_n$ .

$$\dot{n} = H_n \quad (7.7)$$

There are two types of jobs in the economy: research jobs and production jobs. Workers may freely choose which type of job they take, but are assumed to choose the job that gives them the highest income. Although skills are both valuable in research and production, they are appreciated differently. The skills of person  $i$  allow either for a production of  $h_x(i)$  consumption goods or for the invention of  $h_n(i)$  new product designs. Their exact specifications are

$$h_x(i) = ak(i)^\alpha = a\bar{s}^\alpha (i/L)^{\sigma\alpha} \quad (7.8)$$

$$h_n(i) = bk(i)^\beta = b\bar{s}^\beta (i/L)^{\sigma\beta} \quad (7.9)$$

$$\bar{s} \equiv (\sigma + 1)s,$$

where  $a, b > 0$  and  $\beta > \alpha \geq 0$ . The latter condition ensures that the marginal importance of skills is higher for researchers than for production workers. With this setup, the people with relatively high skills will end up in research.<sup>5</sup> What remains to be determined is what level of skills marks the border between production workers and researchers.

The worker that is indifferent between a production job and a research job is indexed  $L_x$ , such that the workers 0 through  $L_x$  produce consumption goods and the workers  $L_x$  through  $L$  invent new products. The worker that is indifferent between production and research, must earn the same income with both kinds of jobs.

$$w_x h_x(L_x) = w_n h_n(L_x) \quad (7.10)$$

Here,  $w_x$  is wage rate per unit of human capital in production, and  $w_n$  is the wage rate per unit of human capital in research. After substitution for  $h_x$  and  $h_n$ , the ratio of the wage rates can be seen to be related to the allocation of labor.

$$\frac{w_x}{w_n} = \frac{b\bar{s}^\beta (L_x/L)^{\sigma\beta}}{a\bar{s}^\alpha (L_x/L)^{\sigma\alpha}} = \frac{b}{a} \bar{s}^{\beta-\alpha} \left( \frac{L_x}{L} \right)^{\sigma(\beta-\alpha)} \quad (7.11)$$

The aggregate amounts of human capital can be found by integration over the appropriate range of the labor force.

$$H_x = \int_0^{L_x} a\bar{s}^\alpha (i/L)^{\sigma\alpha} di = \frac{a\bar{s}^\alpha}{\sigma\alpha + 1} \left( \frac{L_x}{L} \right)^{\sigma\alpha+1} L \quad (7.12)$$

$$H_n = \int_{L_x}^L b\bar{s}^\beta (i/L)^{\sigma\beta} di = \frac{b\bar{s}^\beta}{\sigma\beta + 1} \left( 1 - \left( \frac{L_x}{L} \right)^{\sigma\beta+1} \right) L \quad (7.13)$$

(The constants of integration have been set to zero for convenience.)



The profit value ratio,  $\pi/v$ , can be derived in the same way as with equation 4.68.

$$\frac{\pi}{v} = \frac{(1-\gamma)C}{nw_n} = \frac{1-\gamma}{\gamma} \frac{w_x}{w_n} \frac{H_x}{n} \quad (7.14)$$

Like before, free entry in the research sector implies  $v = w_n$ . After substitution for the ratio of wage rates and human capital employed in production, the profit value ratio becomes

$$\frac{\pi}{v} = \frac{(1-\gamma)b\tilde{s}^\beta}{\gamma(\sigma\alpha+1)} \left(\frac{L_x}{L}\right)^{\sigma\beta+1} \frac{L}{n}. \quad (7.15)$$

Using this last expression the Ramsey rule (see equation 4.72) can be formulated in terms of  $L_x/L$  and  $L/n$ .

$$\hat{C} = \frac{(1-\gamma)b\tilde{s}^\beta}{\gamma(\sigma\alpha+1)} \left(\frac{L_x}{L}\right)^{\sigma\beta+1} \frac{L}{n} + g_L - \bar{\rho} \quad (7.16)$$

Equation 7.6 yields another expression for the growth rate of consumption.

$$\hat{C} = \frac{1-\gamma}{\gamma} \hat{n} + \hat{H}_x \quad (7.17)$$

The growth rates of  $\hat{H}_x$  and  $n$  can be obtained from 7.12 and 7.7 together with 7.13.

$$\hat{H}_x = \alpha g_s + (\sigma\alpha+1) \hat{L}_x - \sigma\alpha g_L \quad (7.18)$$

$$\hat{n} = \frac{H_n}{n} = \frac{b\tilde{s}^\beta}{\sigma\beta+1} \left(1 - \left(\frac{L_x}{L}\right)^{\sigma\beta+1}\right) \frac{L}{n} \quad (7.19)$$

In the first expression,  $g_s$  is the exogenous growth rate of average skills (endogenous skill growth is discussed in appendix 7.A). Substitute for  $\hat{n}$  and  $\hat{H}_x$  to get the growth rate of consumption in terms of  $L_x/L$  and  $L/n$ .

$$\hat{C} = \frac{(1-\gamma)b\tilde{s}^\beta}{\gamma(\sigma\beta+1)} \left(1 - \left(\frac{L_x}{L}\right)^{\sigma\beta+1}\right) \frac{L}{n} + \alpha g_s + (\sigma\alpha+1) \hat{L}_x - \sigma\alpha g_L \quad (7.20)$$

Together, equations 7.16, 7.19, and 7.20 provide sufficient information to study the dynamic behavior of the model.

### 7.2.3 Steady state

Before we proceed with the analysis of the dynamic properties of the model, let us first rephrase the condensed model formed by equations 7.16, 7.19, and 7.20 in order to reduce its complexity. Define  $\Lambda \equiv L_x/L$  and  $\lambda \equiv s^\beta L/n$ . It turns out to be that the steady state of the model coincides with constant values for  $\Lambda$  and  $\lambda$ .<sup>6</sup>

$$\hat{C} = \frac{(1-\gamma)b(\sigma+1)^\beta}{\gamma(\sigma\alpha+1)} \Lambda^{\sigma\beta+1} \lambda + g_L - \bar{\rho} \quad (7.21)$$

$$\beta g_s + g_L - \hat{\lambda} = \frac{b(\sigma+1)^\beta}{\sigma\beta+1} (1 - \Lambda^{\sigma\beta+1}) \lambda \quad (7.22)$$

$$\begin{aligned} \hat{C} = & \frac{(1-\gamma)b(\sigma+1)^\beta}{\gamma(\sigma\beta+1)} (1 - \Lambda^{\sigma\beta+1}) \lambda \\ & + \alpha g_s + (\sigma\alpha+1) \hat{\Lambda} + g_L \end{aligned} \quad (7.23)$$

After substituting out  $\hat{C}$  and solving for  $\hat{\Lambda}$ , we obtain a system of two equations in  $\Lambda$  and  $\lambda$ .

$$\hat{\Lambda} = \frac{(1-\gamma)b(\sigma+1)^\beta}{\gamma(\sigma\alpha+1)} \left( \left( \frac{1}{\sigma\alpha+1} + \frac{1}{\sigma\beta+1} \right) \Lambda^{\sigma\beta+1} - \frac{1}{\sigma\beta+1} \right) \lambda - \frac{\alpha g_s + \bar{\rho}}{\sigma\alpha+1} \quad (7.24)$$

$$\hat{\lambda} = \beta g_s + g_L - \frac{b(\sigma+1)^\beta}{\sigma\beta+1} (1 - \Lambda^{\sigma\beta+1}) \lambda \quad (7.25)$$

The steady state of this system is characterized by a constant share of production workers in the labor force,  $\Lambda$ , and a constant  $\lambda$ . Setting  $\hat{\lambda} = 0$  in 7.25 and  $\hat{\Lambda} = 0$  in 7.24 yields the steady state value of  $\lambda$  as functions of  $\Lambda^*$ , the steady state value of  $\Lambda$  (steady state levels carry a star).

$$\lambda^* = \frac{(\sigma\beta+1)(\beta g_s + g_L)}{b(\sigma+1)^\beta} (1 - \Lambda^{*\sigma\beta+1})^{-1} \quad (7.26)$$

$$\lambda^* = \frac{[(1-\gamma)g_L + \gamma\rho + (\gamma\alpha + (1-\gamma)\beta)g_s](\sigma\alpha+1)}{(1-\gamma)b(\sigma+1)^\beta} \Lambda^{*- \sigma\beta-1} \quad (7.27)$$

The first expression has been used to simplify the second expression. Additionally,  $\bar{\rho}$  has been replaced with  $\rho$  as the profit-value ratio is constant in the steady state (equation 4.68). Equate both expressions for  $\lambda^*$  to get a solution for  $\Lambda^*$  and, after substitution of  $\Lambda^*$ , a solution for  $\lambda^*$  as well.

$$\Lambda^* = \left( \frac{\Theta}{1+\Theta} \right)^{\frac{1}{\sigma\beta+1}} \quad (7.28)$$

$$\lambda^* = \frac{(\sigma\beta+1)(\beta g_s + g_L)}{b(\sigma+1)^\beta} (1 + \Theta) \quad (7.29)$$

$$\Theta \equiv \frac{\sigma\alpha+1}{\sigma\beta+1} \left( 1 + \frac{\gamma(\alpha g_s + \rho)}{(1-\gamma)(\beta g_s + g_L)} \right) \quad (7.30)$$

The steady state growth rate of consumption can be retrieved either by substituting for  $\lambda^*$  in equation 7.21 using 7.27 or by substituting for  $\lambda^*$  in equation 7.23 using 7.26.

$$g_C = \frac{1}{\gamma} g_L + \frac{\gamma\alpha + (1-\gamma)\beta}{\gamma} g_s \quad (7.31)$$

As was to be expected of a semi-endogenous growth model, the growth rate of consumption in the steady state depends on the growth rate of the population. A novel feature here is that consumption growth also depends on the growth rate of average skills. A substantial difference with the standard semi-endogenous growth model presented in section 4.1 is that steady state economic growth is also feasible in the absence of population growth. However, as both population growth and skill growth are exogenous, the label 'semi-endogenous' is still appropriate.

The solution for  $g_C$  in 7.31 could also have been found using a shortcut. The steady state growth rates of  $\hat{H}_x$  and  $\hat{n}$  are given by

$$g_{H_x} = \alpha g_s + g_L \quad (7.32)$$

$$g_n = \beta g_s + g_L. \quad (7.33)$$

Applying these growth rates to equation 7.17 immediately yields the steady state growth rate of consumption. Above expressions clearly illustrate that growth in average skills raises both the productivity of production workers and researchers. By doing so, advances in education affect economic growth in much the same way as population growth does.

An interesting detail is that  $\Lambda^*$  is smaller than one if population growth is zero but skill growth is positive (see equation 7.28). Even when the population is fixed, researchers are employed and new products are introduced to the market. This is why the steady state growth rate of consumption is higher than the rate of productivity growth in the production sector as long as  $g_s > 0$  (remember  $\beta > \alpha$ ). This feature adds a flavor of neo-endogenous growth to the model.<sup>7</sup>

### 7.2.4 Schooling inequality

We have seen that a change in the parameter  $\sigma$  alters the shape of the skill distribution without affecting its mean. The change in the shape of the distribution is such that a rise in  $\sigma$  widens the gap between the minimum and the maximum skill level, while simultaneously increasing the variance of the skill distribution. A reduction in  $\sigma$  has the reverse effect. This property of  $\sigma$  makes it a suitable measure of schooling inequality – at least within the context of the model.

The steady state values of  $\lambda$  and  $\Lambda$  have been derived in the previous section. Equations 7.28, 7.29, and 7.30 show that the steady state values are dependent on  $\sigma$ : schooling inequality matters for the amount of research being done as well as the number of product types available for consumption. The fact that  $\sigma$  occurs several times in each of these equations indicates that the impact of a change in  $\sigma$  is quite complex. Below we will analyze the effects of a change in  $\sigma$  on the steady state in three steps.

In the first step it will be shown how  $\sigma$  affects the kind of job – production or research – that is preferred by worker  $L_x$ , while keeping the wage rates constant. This effect of  $\sigma$  on the labor market is labelled the ‘skill’ effect. With the second step it is shown how the wage rates will adjust after the ‘skill’ effect has taken place. The adjustment of the wage rates naturally causes workers to reconsider their job choice. This second effect is labelled the ‘wage’ effect. The third and last step involves the effect of  $\sigma$  on product variety given that both the ‘skill’ and the ‘wage’ effect have taken place and the proportion of production workers has reached the steady state.

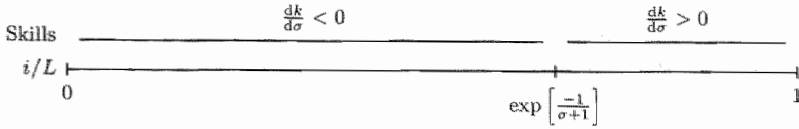
The three steps do not reflect the transitional dynamics of the model and are only used to make the comparative static effects of a change in  $\sigma$  more insightful.

The first step starts by analyzing how the skills of an individual worker are affected by inequality, after this the effects on his income are discussed. A change in the shape of the skill distribution may have a positive or a negative effect on the skills of person  $i$ , depending on his ranking. The skills of person  $i$  will increase in response to a rise in  $\sigma$  if the following condition holds:

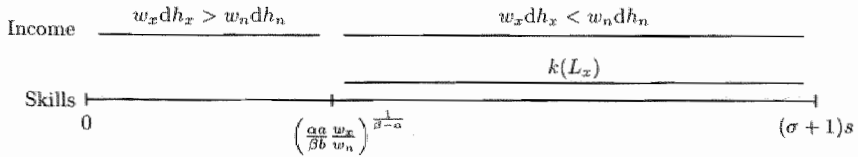
$$i/L > \exp \left[ \frac{-1}{\sigma + 1} \right] \quad (7.34)$$

This condition is obtained by differentiating equation 7.1 with respect to  $\sigma$ . A graphical representation is given in figure 7.2.

A change in skills affects the amount of human capital a worker can supply to production or research. A worker will be more inclined to do research if  $w_x dh_x < w_n dh_n$  and he will be more inclined to take a production job if  $w_x dh_x > w_n dh_n$



**Figure 7.2.** Change in the skills of worker  $i$  due to a change in inequality



**Figure 7.3.** Effect of a rise in skills on income (constant wages) and skill domain of worker  $L_x$

(we keep wages fixed for the moment). The change in the relative attractiveness of the jobs can be found by differentiating equations 7.8 and 7.9 with respect to  $k(i)$ . The condition below marks the skill level at which  $w_x dh_x = w_n dh_n$ .

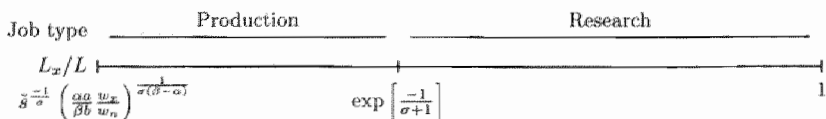
$$k(i) = \left( \frac{\alpha a}{\beta b} \frac{w_x}{w_n} \right)^{\frac{1}{\beta-\alpha}} \quad (7.35)$$

A skill level that exceeds this value will encourage workers to do research. If the skills of a worker are lower than this value, a marginal increase in skills will raise the attractiveness of a production job. Low-skilled people benefit from a higher skill level because it makes them more productive in their current occupation. Their productivity as a researcher remains very low, causing their wage gap between production and research to widen and not to reduce. The reverse applies to high skilled production workers. A rise in their skills will reduce the difference between their current income and the income that they would earn in research. The first line in figure 7.3 shows how the attractiveness of a job depends on the skills of the worker.

In general, a change skills can either raise or lower the attractiveness of a job in research, depending on the skill level of the worker. However, there is only one worker that might actually switch jobs: worker  $L_x$ . Can we be more specific about the incentives faced by  $L_x$ ? Fortunately, we can. Use equation 7.11 to solve for  $k(L_x)$  as a function of the ratio of wages and compare the outcome with the skill level for which  $w_x dh_x = w_n dh_n$ .

$$k(L_x) = \left( \frac{a}{b} \frac{w_x}{w_n} \right)^{\frac{1}{\beta-\alpha}} > \left( \frac{\alpha a}{\beta b} \frac{w_x}{w_n} \right)^{\frac{1}{\beta-\alpha}} \quad (7.36)$$

This leaves us with the clean result that if the skills of worker  $L_x$  increase, then he will choose to be a researcher; if they decrease, he will choose a job in production. This is the 'skill' effect: after a change in  $\sigma$ , worker  $L_x$  can improve his income by switching jobs because his skill level has changed. Figure 7.4 contains a graphical representation of this result. The domain labeled 'Production' is where worker  $L_x$  chooses a production job; the domain labeled 'Research' is where he chooses to become a researcher.



**Figure 7.4.** Type of job chosen by worker  $L_x$  in response to a rise in inequality (constant wages)

So far, we have analyzed the effects of a change in inequality keeping wage rates constant. However, a change in inequality is unlikely to leave wage rates unaffected. The underlying reason is that a change in inequality will affect both types of human capital. First, there is a direct effect through the presence of  $\sigma$  in equations 7.8 and 7.9. Second,  $\sigma$  is important for the job choice of worker  $L_x$ . If  $H_x$  and  $H_n$  change, then – in general – there will be over- or under-investment in research. When this happens a change in wages is required to bring the economy back to the steady state.

The complexity of the model makes it impractical to discuss the impact of a change in  $\sigma$  on the wages through its effect on  $H_x$  and  $H_n$ . In stead, I will discuss the change in the wage rates using the ratio of wage bills as this is mathematically more convenient. An expression for the ratio of the wage bills can be derived using the labor market skill equation (7.11) in combination with the definitions of human capital (7.8, 7.9).

$$\frac{w_x H_x}{w_n H_n} = \frac{\sigma\beta + 1}{\sigma\alpha + 1} \frac{1}{\left(\left(\frac{L_x}{L}\right)^{-\sigma\beta-1} - 1\right)} \quad (7.37)$$

The steady state value of the wage bill ratio follows from substituting  $L_x/L$  with  $\Lambda^*$ , which is given by 7.28.

$$\left(\frac{w_x H_x}{w_n H_n}\right)^* = 1 + \frac{\gamma(\alpha g_s + \rho)}{(1 - \gamma)(\beta g_s + g_L)} \quad (7.38)$$

The convenient property of the wage bill ratio is that it is independent of  $\sigma$  in the steady state. A change in  $\sigma$  will therefore only have temporary effects on the wage bill ratio.

If we differentiate the log of the wage bill ratio in equation 7.37 with respect to  $\sigma$  while keeping  $L_x/L$  constant, we find that the sign of this derivative depends on  $L_x/L$ .

$$\frac{d}{d\sigma} \ln \left( \frac{w_x H_x}{w_n H_n} \right) = \frac{\beta}{\sigma\beta + 1} - \frac{\alpha}{\sigma\alpha + 1} + \frac{\beta \ln \left[ \frac{L_x}{L} \right]}{\left(\frac{L_x}{L}\right)^{-\sigma\beta-1} - 1} \quad (7.39)$$

The precise value of  $L_x/L$  for which  $\frac{d}{d\sigma} \ln \left( \frac{w_x H_x}{w_n H_n} \right) = 0$  can be found by numerically solving the following equation:

$$\frac{\beta}{\sigma\beta + 1} - \frac{\alpha}{\sigma\alpha + 1} = \frac{-\beta \ln \left[ \frac{L_x}{L} \right]}{\left(\frac{L_x}{L}\right)^{-\sigma\beta-1} - 1}. \quad (7.40)$$

I will label the solution to this equality  $\Lambda^\#$ . There will be at most one solution as both  $\beta \ln \left[ \frac{L_x}{L} \right]$  and  $\left(\left(\frac{L_x}{L}\right)^{-\sigma\beta-1} - 1\right)^{-1}$  are monotonically increasing in  $L_x/L$ .

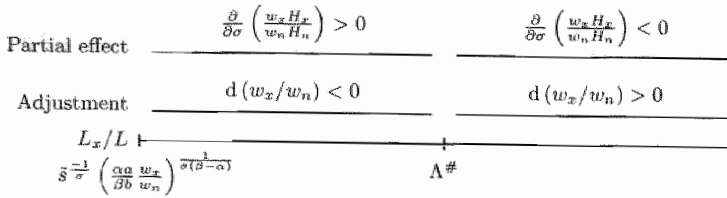


Figure 7.5. Adjustment of wage ratio to steady state

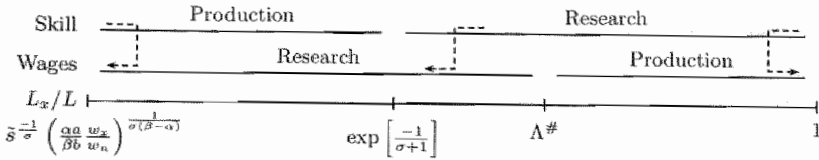


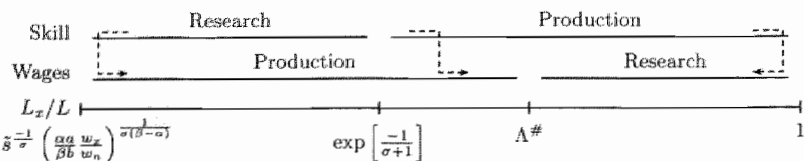
Figure 7.6. Job choice by worker  $L_x$  in response to a rise in inequality: skill and wage effects ( $\exp \left[ \frac{-1}{\sigma+1} \right] < \Lambda^\#$ )

A higher  $\sigma$  causes the wage bill ratio to rise above its steady state value if the proportion of production workers is lower than  $\Lambda^\#$ ; the wage bill ratio will decline if  $L_x/L > \Lambda^\#$ . The first line in figure 7.5 refers to this partial effect of a change in  $\sigma$  on the wage bill ratio.

When the wage bill ratio deviates from its steady state value, an adjustment on the labor market needs to take place to reach the steady state again. Suppose a rise in inequality leads to an increase in the wage bill ratio, then a return to the steady state requires a decrease in the wage bill for production relative to that of research. This can only be accomplished by a drop in  $w_x$  relative to  $w_n$ . Alternatively, if  $L_x/L > \Lambda^\#$ , then the wage rate for researchers is too high relative to the wage rate for production workers. The second line in figure 7.5 shows how wages adjust to a change in  $\sigma$ .

Above we have first established the effect of a change in  $\sigma$  on job choice keeping wage rates constant. Second, we have established the effect of a change in  $\sigma$  on the wage rates assuming that workers choose their jobs optimally. Combining the two effects allows us to analyze the overall comparative static effects of a rise in  $\sigma$ . As we rely on equation 7.11 for the analysis of wage adjustment, we will have to start with analyzing the 'skill' effect before we can turn to the 'wage' effect. The overall comparative static effects are summarized in figure 7.6. The 'skill' effect is shown on the first line, which is identical to figure 7.4. The second line shows the 'wage' effect assuming that the 'skill' effect has already taken place.

The three dashed arrows represent three scenario's for arriving at a new steady state when inequality increases. The leftmost arrow shows the response of worker  $L_x$  if a large part of the workforce is employed in research. First, worker  $L_x$  finds out that his skill level is lower, which induces him to take a production job. Second, the new worker  $L_x$  is confronted with a decline in  $w_x/w_n$  causing him to become a researcher. Which of the two effects is dominant depends on the precise parameter values. The rightmost arrow describes the opposite situation. Worker  $L_x$  experiences a rise in skills and decides to do research. The second worker  $L_x$  sees a rise in  $w_x/w_n$  and takes a production job.



**Figure 7.7.** Job choice by worker  $L_x$  in response to a decline in inequality: skill and wage effects ( $\exp \left[ \frac{-1}{\sigma+1} \right] < \Lambda^{\#}$ )

The middle arrow shows a situation in which the ‘skill’ and the ‘wage’ effects work in the same direction. If  $\exp \left[ \frac{-1}{\sigma+1} \right] < L_x/L < \Lambda^{\#}$  like in figure 7.6, higher inequality will cause the proportion of researchers in the workforce to increase. If  $\Lambda^{\#} < L_x/L < \exp \left[ \frac{-1}{\sigma+1} \right]$ , then the shift will be towards production.

In view of the empirical trend in OECD countries towards a more equal distribution of educational attainment, it is interesting to briefly discuss the effects of a decline in  $\sigma$ . Figure 7.7 summarizes these effects. Not surprisingly, the figure is a ‘mirror image’ of figure 7.6 as all effects work exactly in the opposite direction. Another empirical trend – one that has provided the empirical basis of the Jones critique summarized in section 4.1 – is the rising share of researchers in the workforce (figure 3.9). Assuming it is appropriate to consider the proportion of researchers in OECD countries to be ‘small’ (meaning  $L_x/L > \Lambda^{\#}$ ), the model is able to explain how a reduction in schooling inequality can lead to a rising proportion of researchers.<sup>8</sup> In particular, the comparative static analysis demonstrates that a continuously rising proportion of researchers can be a property of an economy that is in the steady state. A growing proportion of researchers need not be a transitional effect if schooling inequality is also continuously declining. Of course, this can only be true if the ‘skill’ effect is smaller than the ‘wage’ effect.

On the rightmost part of the domain sketched in figure 7.7 the model yields the prediction that a more equal skill distribution raises the wage rate for researcher relative to that of production workers. This result comes very close to what Acemoglu (1998) calls the ‘strong induced-bias hypothesis’: “... directed technical change can make the long-run relative demand curve [for skilled labor] slope up.” (p. 783). Of course, the model presented here differs in one important respect from Acemoglu’s approach: it does not involve skill-biased technical change!

Having analyzed the effects of schooling inequality on the proportion of production workers, what remains to be done is the analysis of the effects on product variety. Equation 7.26 shows that  $\lambda^*$  depends on  $\sigma$  both directly and indirectly through  $\Lambda^*$ . The direct effect is due to a change in the average productivity of researchers and is given by the partial derivative of  $\lambda^*$  with respect to  $\sigma$ . Unfortunately, the sign of the partial derivative cannot be found analytically.

$$\frac{\partial \ln \lambda^*}{\partial \sigma} = \left( \frac{\beta}{\sigma\beta + 1} - \frac{\beta}{\sigma + 1} + \frac{\beta \ln [\Lambda^*]}{\Lambda^{*- \sigma\beta - 1} - 1} \right) \lambda^* \quad (7.41)$$

The solution for  $\partial \ln (\lambda^*) / \partial \sigma = 0$  is labeled  $\Lambda^{*\#}$  and can be found by solving the following equation numerically:<sup>9</sup>

$$\frac{\beta}{\sigma\beta + 1} - \frac{\beta}{\sigma + 1} = \frac{-\beta \ln [\Lambda^*]}{\Lambda^{*- \sigma\beta - 1} - 1} \quad (7.42)$$

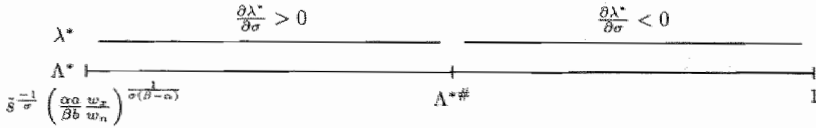


Figure 7.8. Product variety in the steady state

Figure 7.8 shows the domain of  $\Lambda^*$  for which a rise in  $\sigma$  has a positive direct effect on product variety per capita ( $\partial \lambda^* / \partial \sigma < 0$ ) and the domain for which the direct effect is negative ( $\partial \lambda^* / \partial \sigma > 0$ ).

The indirect effect depends on how  $\Lambda^*$  is affected by  $\sigma$ . We have seen above that this relationship is quite complex. Once we know the change in  $\Lambda^*$ , however, remaining part of the analysis is straightforward as  $\partial \lambda^* / \partial \Lambda^*$  is always positive. This follows directly from the fact that a larger number of researchers invent a greater number of new products as equation 7.26 is just a reformulation of  $\dot{n} = H_n$ .

We have noted before that schooling inequality has declined in the OECD, while the proportion of researchers has risen. Given these facts, the model predicts that the total effect of the decline in schooling inequality on product variety per capita has probably been positive, provided that the direct effect on  $\lambda^*$  is not too strong if positive.

### 7.2.5 Summary

The analysis of this section demonstrated how economic growth is affected by growth in the level of education and changes in the distribution of skills. Advances in schooling induce economic growth, not only in the trivial way of raising a worker's productivity, but also because it stimulates research in the same way as population growth does. Because of the latter effect, consumption per capita grows at a higher rate than the productivity of production workers does.

The effects of a change in the shape of the distribution of skills on product variety are ambiguous. Although the specification of the model outlined above is still fairly simple, it has proved to be difficult to draw some general conclusions. A closed form solution has been found for the steady state, but how product variety is affected by a change in schooling inequality depends entirely on parameter settings. Given certain parameter settings, the model does provide an explanation for some empirical trends. The model shows that a reduction in schooling inequality might raise the proportion of researchers in the workforce, while the wage rate for researchers rises relative to that for production workers.

## 7.3 Imperfect improvements

### 7.3.1 Introduction

The class of models outlined in chapter 4 had two kinds of research activities. Researchers were either employed to invent new products or to design improved versions of existing products. In practise, the supposed dichotomy between variety expansion and quality improvements does not seem to work very well. Many new products are not equivalent to existing product types in terms of usefulness, but



at the same time they cannot be considered to be a new version of an existing product that renders the older version obsolete. For the majority of innovations, it seems to be more appropriate to think of them as being 'imperfect improvements' over existing types. As an example, consider the case of the central processing unit (CPU), the core of a personal computer. The introduction of a new type of CPU is commonly thought of as the perfect example of a quality improving innovation. Although in most cases new CPUs are indeed simply faster than older types, there have also been innovations in different directions. One such innovation is the introduction of CPUs with a variable processing speed. These so-called 'mobile' processors enable laptops to work longer on the same battery pack, supplying only the computing power that is actually needed. The curious property of CPUs of this type is that they are slower than ordinary processors. Hence, 'mobile' CPUs can be considered to be imperfect improvements over ordinary CPUs. Another example in this direction is Intel's Centrino chipset that facilitates wireless networking, thereby enabling smaller laptops that consume less energy. The Centrino chipset, in turn, is an imperfect improvement over ordinary chipsets with a 'mobile' processor: the CPU that comes with a Centrino chipset operates at a frequency that is significantly lower than that of a normal 'mobile' processor.

Numerous other examples of imperfect improvements can be given. The invention of the aeroplane did not end the production of ships or trains. Nuclear power plants did not drive out coal-fired power plants – except in France. Computers did not even make pocket calculators obsolete. In a similar vein, even when a new product is undoubtedly better than previous versions, it may still be considered an imperfect improvement if its production costs are higher. When this is the case, both types can coexist in the market provided that consumers have different incomes or preferences.

There is a fundamental reason to expect quality improvements to be imperfect. Often, a product has at least one significant drawback, e.g. high energy consumption. Researchers will try to eliminate this drawback or, if that is not possible, to decrease its magnitude. Hereto, they will have to make changes in the design of the product, which might confront the researcher with a trade-off between different desirable characteristics of the product. In the case of CPUs, higher speed comes at the cost of higher energy consumption and heat generation.

A study by Bils and Klenow (2001) reflects the importance of imperfectly improving innovations. They used data from the United States' Bureau of Labor Statistics (BLS) to construct a measure for the 'dynamism' of product groups. According to their definition a dynamic product group is characterized by a high item-substitution rate, i.e. the rate at which products that are discontinued at a particular outlet are replaced with 'dissimilar' products. As such, the item substitution rate reflects the intensity with which new goods depress the sales of older goods. Bils and Klenow's main finding is that consumer spending has shifted away from product groups with little variety change towards more dynamic product groups, even though they control for income effects.

This section presents a simple model of semi-endogenous growth that applies to the grey area between variety expansion and quality improvements. It allows for intertemporal knowledge spillovers of a variable degree, but is not plagued by the knife-edge configurations that play such an important role in chapter 4. A special feature is that semi-endogenous growth is also feasible when spillovers are negative but not too strong.

The upshot of the model is that quality improvements may cause the variety of products to expand, but that this kind of variety expansion is not as valuable to

the consumer as the traditional kind where all product types are equally important to the consumer. This implies that the 'effective' variety of consumption goods is smaller than the number of product types would suggest.

Spillovers enter the model in a different way than is customary. In the generalized growth model of chapter 4 knowledge spillovers affect the arrival rates of new products and quality improvements (see equation 4.34). A greater number of product types or a higher quality level is assumed to stimulate the number of inventions per researcher. Also another specification for spillovers has been used. The early models by Grossman and Helpman (1991a) and Aghion and Howitt (1992), or GHAA, rely on the mechanism that the current quality level affects the size of the upcoming quality improvement. The particular specification they use is  $q(ij) = a^j$ , where  $j$  is the number of quality improvements that have been achieved for good  $i$ . The arrival of such an improvement is stochastic but depends on the number of researchers that is hired. Suppose nevertheless that  $j$  is a continuous and differentiable function of time, then the rate of quality change is given by  $\dot{q} = \ln(a) q(ij) \phi l_q$ , where  $dj/dt$  has been replaced with the arrival rate,  $\phi$ , multiplied by the number of researchers,  $l_q$ . The insight of this exercise is that, unless  $\phi$  is negatively related to  $q$ , GHAA spillovers imply endogenous growth.

The model presented here does not have spillovers that (directly) affect arrival rates like in chapter 4. In stead, the perspective of GHAA that new products are supposed to be better than old products is adopted. However, the model differs from the GHAA approach in two respects. First, spillovers, no matter how strong, do not lead to endogenous growth but they do affect the growth rate of the economy when the population is growing. Second, a new product does not replace an existing product completely, but it does depress the sales of existing products – it is an imperfect improvement. The model incorporates a mild form of 'creative destruction' as the introduction of a new product might cause the profitability of older products to decline over time.

Although the number of product types does not influence the number of patents a single researcher delivers per unit of time, it does affect the quality of those patents. Therefore, a larger stock of patents raises the incentives for developing a new patent – provided that spillovers are positive. In this sense, the structure of spillovers adopted here is similar to the standard spillover structure (e.g.  $\dot{n} = n\beta L_R$ ).

Spillovers are modeled in a similar way by Van Zon, Fortune, and Kronenberg (2003, pp. 12-13). They let every new peripheral to a general purpose technology be less productive than its predecessors, implying decreasing returns to scale for applied research. Although negative spillovers are also possible in the model presented below, I will focus on positive rather than negative spillovers.

Analysis of patent citations has shown that knowledge spillovers tend to be both internal to industries and geographically concentrated (Jaffe 1986, Maurseth and Verspagen 2002). In the model of this section no explicit separation between industries is assumed. Knowledge effectively spills over from the latest product to the next. A possible extension of the model would be to include the effects of (technological) distance in the spirit of Peretto and Smulders (2002). However, this is not attempted here.

The model is introduced and its steady state is derived in subsection 7.3.2. The main insights of the model are discussed in subsection 7.3.3.

### 7.3.2 The model

The workforce,  $L$ , is subdivided over two activities: the production of consumer goods and the invention of new products. Total employment in the production and research sector are  $L_x$  and  $l_n$ , respectively. Consumers have the usual CES preferences but because each new product is more important than its predecessors products are no longer symmetric. In fact, the quality of a product now depends on its index, assuming that older product types have a lower index than newer types. The quality of product  $i \in [0, n]$  is given by  $i^\nu$ , where  $\nu > \gamma - 1$  is the degree of superiority of a new good over the previously introduced goods. The aggregate consumption index,  $C$ , has the following form:

$$C = \left( \int_0^n i^\nu x(i)^\gamma di \right)^{\frac{1}{\gamma}} \quad (7.43)$$

Standard formulations can be obtained by varying the parameter settings. if  $\nu = 0$  we get the standard love-of-variety model; if  $\gamma = 1$  each new product is a perfect improvement over the previous version just as in the standard quality-ladder model. The consumer goods are produced in the familiar way,  $x(i) = bl(i)$ , where  $l(i)$  is the amount of labor used and  $b$  is a parameter. Maximization of  $C$  subject to a budget constraint yields the inverse demand curve.<sup>10</sup>

$$p(i) = \lambda^{-1} C^{1-\gamma} i^\nu x(i)^{\gamma-1} \quad (7.44)$$

Here,  $\lambda$  is the Lagrange multiplier.

Profit maximization by the producers of consumer goods results in prices being a markup over the wage rate,  $w$ .

$$p(i) = \frac{w}{\gamma b} \equiv p \quad (7.45)$$

Like we have seen in section 4.2, every good is being sold at the same price despite heterogeneity in quality. With this result in hand, the distribution of income from production over wages and profits can be determined. As  $\int_0^n px(i) di = C$ , replacing  $p$  with  $w/(\gamma b)$  returns  $wL_x = \gamma C$ , and, consequently,  $n\pi = (1 - \gamma)C$ .

The demand for product  $i$  can be found by equating both expressions for the price.

$$x(i) = \left( \frac{w\lambda}{\gamma b i^\nu} \right)^{\frac{1}{\gamma-1}} C \quad (7.46)$$

The sales of a product depend on its index. The chosen specification for the quality level ensures that the amount consumed is higher when a product is newer. A refined solution for  $x(i)$  can be found by computing the share of the workforce employed in sector  $i$  and by substituting for  $l(i)$ .

$$\frac{l(i)}{L_x} = \frac{\left( \frac{w\lambda}{\gamma b i^\nu} \right)^{\frac{1}{\gamma-1}} \frac{C}{b}}{\int_0^n \left( \frac{w\lambda}{\gamma b (i')^\nu} \right)^{\frac{1}{\gamma-1}} \frac{C}{b} di'} = \frac{i^{\frac{\nu}{1-\gamma}}}{\int_0^n i'^{\frac{\nu}{1-\gamma}} di'} \quad (7.47)$$

$$x(i) = \frac{i^{\frac{\nu}{1-\gamma}}}{\int_0^n i'^{\frac{\nu}{1-\gamma}} di'} b L_x \quad (7.48)$$

Insert this last expression into equation 7.43 and integrate to get the consumption index as a function of the number of different products and the number of production workers.

$$C = \left( \int_0^n i^{\frac{\nu}{1-\gamma}} di \right)^{\frac{1-\gamma}{\gamma}} bL_x = \left( \frac{1-\gamma}{\nu+1-\gamma} \right)^{\frac{1-\gamma}{\gamma}} n^{\frac{\nu+1-\gamma}{\gamma}} bL_x \quad (7.49)$$

The constant of integration has been set to zero. The growth rate of consumption depends on the growth of product variety and the growth of the workforce.

$$\hat{C} = \frac{\nu+1-\gamma}{\gamma} \hat{n} + \hat{L}_x \quad (7.50)$$

The profits of firm  $i$  equal the profit margin per product times the quantity being sold.

$$\pi(i) = \left( p - \frac{w}{b} \right) x(i) = \frac{1-\gamma}{\gamma b} w x(i) \quad (7.51)$$

With prices being constant across products, heterogeneity in profits is purely the result of heterogeneity in the quantities sold to consumers. Substitute for  $x(i)$  using equation 7.48.

$$\pi(i) = \frac{\nu+1-\gamma}{\gamma} i^{\frac{\nu}{1-\gamma}} n^{\frac{\gamma-1-\nu}{1-\gamma}} wL_x \quad (7.52)$$

Replace  $wL_x$  with  $\gamma C$  and substitute for  $C$  using equation 7.49 to get an expression for the profits of firm  $i$  in terms of  $i$ ,  $n$ , and  $L_x$ .

$$\pi(i) = (\nu+1-\gamma) b \left( \frac{1-\gamma}{\nu+1-\gamma} \right)^{\frac{1-\gamma}{\gamma}} i^{\frac{\nu}{1-\gamma}} n^{\frac{\gamma-1-\nu}{1-\gamma} + \frac{\nu+1-\gamma}{\gamma}} L_x \quad (7.53)$$

Profits can be seen to depend on the quality of the product,  $i^\nu$ , the number of competitors,  $n^{\frac{\gamma-1-\nu}{1-\gamma}}$ , the standard of living,  $n^{\frac{\nu+1-\gamma}{\gamma}}$ , and the size of the market as indicated by  $L_x$ . The standard of living effect and the market size effect enter the profit function via aggregate consumption (equation 7.49). These effect will be discussed in more detail in the next subsection.

The invention of a new product requires the work of one researcher, such that  $\hat{n} = l_n$ . As usual, there are no barriers to entry in the research sector, causing the price of a design to equal the wage rate,  $v = w$ . For a newly invented product, the profit value ratio is therefore a direct result of 7.52.

$$\frac{\pi(n)}{v} = \frac{\nu+1-\gamma}{\gamma} \frac{L_x}{n} \quad (7.54)$$

The profit value ratio for new firms can be inserted in the Ramsey rule, equation 4.60.<sup>11</sup>

$$\hat{C} = g_L + \frac{\nu+1-\gamma}{\gamma} \frac{L_x}{n} - \tilde{\rho} \quad (7.55)$$

The term  $g_L$ , the growth rate of the population, stems from  $C = Lc$ .

The dynamic behavior of the model described above is that of a simple semi-endogenous growth model, like the one described in section 4.1. The steady state growth rate of product variety,  $g_n$ , is equal to the growth rate of the workforce,  $g_L$ . By equation 7.50, the steady state growth rate of per capita consumption is

$$g_C - g_L = \frac{\nu+1-\gamma}{\gamma} g_L. \quad (7.56)$$

Per capita consumption growth directly depends on the spillover parameter  $\nu$ . Independent of the degree of (positive) spillovers, economic growth is only feasible when the population is growing. A different formulation for spillovers can change this result. In particular, it is shown in appendix 7.B that replacing  $i^\nu$  with  $e^{\nu i}$  leads to endogenous growth.

As  $g_n = g_L$ , the profit-value ratio is the same for each new firm. This ensures that  $\bar{p}$  in equation 7.55 is equal to  $\rho$  in the steady state. The number of product types per capita can be found using 7.55.

$$\frac{n}{L} = \frac{\nu + 1 - \gamma}{\gamma \rho} \quad (7.57)$$

Like the growth rate of consumption, the number of products per capita is positively related to the spillover parameter. This result combined with  $g_n = l_n/n$  yields the share of the workforce employed as a researcher.

$$\frac{l_n}{L} = \frac{\nu + 1 - \gamma}{\gamma \rho} g_L \quad (7.58)$$

Besides per capita product variety, also the proportion of researchers depends positively on the intensity of the knowledge spillovers.

### 7.3.3 Discussion

The distance in quality between some firm  $i$  and the market leader, firm  $n$ , grows as new products are introduced. This has several consequences for individual firms. First of all, as can be seen from 7.48,  $x(i)$  is positively related to  $i$ . Firms that sell newer products are larger. Second, with the price being the same for all products, larger firms automatically also have higher profits, witness equation 7.53. The same equation allows for the derivation of the growth rate of profits for individual firms.

$$\hat{\pi}(i) = (\nu + 1 - \gamma) \left( \frac{1}{\gamma} - \frac{1}{1 - \gamma} \right) \hat{n} + \hat{L}_x \quad (7.59)$$

Clearly, whether the growth rate of profits is positively or negatively affected by the growth of product variety depends on whether  $\gamma$  is smaller or larger than a half. The two counteracting forces that underlie this phenomenon are a standard of living effect and a competition effect. First, the introduction of new goods improves the standard of living. Consequently, the price of a single unit of some good rises as prices are expressed in terms of the consumption index. In the expression above, the standard of living effect is captured by  $(\nu + 1 - \gamma) \hat{n}/\gamma$  and follows from equations 7.49 and 7.53.

The standard of living effect is opposed by a competition effect. The creation of a new product leads to the partial destruction of the profits of firms producing older types of goods. Formulated dynamically, the higher the rate of firms entering the market, the lower will be the growth rate of profits for incumbent firms. The size of the competition effect in the expression above is  $(\nu + 1 - \gamma) \hat{n}/(\gamma - 1)$ . When  $\gamma < 1/2$ , the substitutability between products is low and there is less competition between firms causing the standards of living effect to dominate the competition effect. When  $\gamma > 1/2$ , competition is strong enough to curb the standard of living effect, such that the introduction of a new product lowers the profits associated with existing products, given the market size effect induced by population growth.

The cumulative effect of knowledge spillovers makes the profits of the market leader grow at a higher rate than those of individual firms. Having  $i = n$  in equation 7.53 returns the growth rate of the profits of the  $n$ th firm.

$$\hat{\pi}(n) = \left[ (\nu + 1 - \gamma) \left( \frac{1}{\gamma} - \frac{1}{1 - \gamma} \right) + \frac{\nu}{1 - \gamma} \right] \hat{n} + \hat{L}_x \quad (7.60)$$

A higher value of  $\nu$  raises the initial profits of a starting firm relative to the initial profits of its predecessors.

The setup of the model also allows for 'negative spillovers', meaning that each new product has a lower quality than the existing products have. Technically, spillovers are negative when  $\nu < 0$ . A rationale for a negative relation between the novelty of a design and its quality can be that early inventions are being done first exactly because they are the most useful. Early inventions are like low-hanging fruit: the first inventions are very useful, but it becomes increasingly difficult to make an improvement. Although this line of reasoning might be applicable to specific research areas it seems difficult to defend negative 'spillovers' in general, witness the importance of relatively recent inventions in everyday life (computers for example). From equation 7.52 follows that if  $\nu \leq \gamma - 1$ , negative spillovers are too strong for research to be attractive, and, as a consequence, the number of products will not change.

One important difference with the models presented earlier is that not all product types are equivalent. If consumers would have to abstain from consuming one type of product, they would obviously choose to stop consuming the oldest one. The heterogeneity of products thus drives a wedge between the number of product types and the variety of products as experienced by consumers. To give an example, if there are hundred different products but a consumer spends ninety percent of his budget on one good, say beer, then the variety of available goods is not very important to this consumer. Alternatively, another consumer might spend a maximum of ten percent of his budget on a single type of good. For this second person product variety is much more relevant than for the first. Effective product variety differs from the plain number of product types in that it takes into account the relative importance of the individual product types.

It is argued in chapter 5 that the antilog of Shannon's entropy is an appropriate measure for 'effective' product variety (see also chapter 6). The formula of the index for a continuum of products,  $V$ , is stated below; its properties and derivation can be found in chapter 5.

$$V(N) = \exp \left[ \int_0^n -\omega(i) \ln \omega(i) di \right] \quad (7.61)$$

$$\omega(i) \equiv px(i)/C \quad (7.62)$$

The weight  $\omega(i)$  is the share of product  $i$  in the value of the set  $N$  of all products. This measure can be interpreted as a 'weighted number of products'. If all products have the same weight, product variety is maximal and  $V(N) = n$ . In case some products have a higher weight than others  $V(N) < n$ .

With the help of equation 7.48, the weight of product  $i$  can be expressed in terms of  $i$  and  $n$ .

$$\omega(i) = \frac{px(i)}{\int px(i') di'} = \Phi i^{\frac{\nu}{1-\gamma}}, \quad \Phi \equiv \frac{\nu+1-\gamma}{1-\gamma} n^{\frac{\gamma-1-\nu}{1-\gamma}} \quad (7.63)$$

The variety index can be found by solving the integral in 7.61.

$$\ln V(N) = - \int_0^n \Phi_i^{\frac{\nu}{1-\gamma}} \ln \left[ \Phi_i^{\frac{\nu}{1-\gamma}} \right] di \quad (7.64)$$

$$= - \frac{(1-\gamma)\Phi}{(1-\gamma+\nu)^2} n^{\frac{1-\gamma+\nu}{1-\gamma}} \left( -\nu + (1-\gamma+\nu) \ln \left[ \Phi n^{\frac{\nu}{1-\gamma}} \right] \right) \quad (7.65)$$

Details on the integration procedure can be found in appendix 7.C. Next, substitute for  $\Phi$  and exponentiate the result.

$$V(n) = \exp \left[ \frac{\nu}{1-\gamma+\nu} \right] \frac{1-\gamma}{1-\gamma+\nu} n \quad (7.66)$$

The variety index  $V$  turns out to be a linear function of the number of product types. It can easily be verified that if all products have the same weight because spillovers are absent, i.e.  $\nu = 0$ , then the variety index simply equals the number of products,  $V = n$ . Differentiation of  $V$  with respect to  $\nu$  reveals that, for any given number of products, higher positive spillovers imply a lower effective product variety.

$$\frac{dV(n)}{d\nu} = - \exp \left[ \frac{\nu}{1-\gamma+\nu} \right] \frac{\nu(1-\gamma)}{1-\gamma+\nu} n \quad (7.67)$$

A high degree of spillovers causes new products to have a much larger weight than existing products. As consumers do not consider older products to be important, the product variety experienced by the consumer is lower than the number of products would suggest.

When spillovers are negative, the reverse applies: for  $\nu < 0$ , a higher value of  $\nu$  increases effective product variety. Negative spillovers cause the weights of new products to be lower than those of old products. For this reason weaker negative spillovers ( $\nu$  closer to zero) cause product variety to be higher. It is not the sign of the spillovers but their absolute magnitude that matters for effective product variety.

## 7.4 Conclusion

Two models have been presented that provide new insights in how education and knowledge spillovers may affect economic growth and product variety. The first model shows how growth in educational attainment can induce economic growth. Advances in schooling not only raise productivity, but they also stimulate research in the same way as population growth does. Because of the latter effect, consumption per capita grows at a higher rate than the productivity of production workers does. The model also demonstrates how the structural decline in schooling inequality might lead to a structurally rising proportion of researchers in the workforce.

The second model presented in this chapter covers the spectrum between quality-improving and variety-expanding innovations. It provides an answer to the question why the sales of existing products experience pressure from newly introduced products, while, at the same time, this pressure is not large enough to cause these older products to disappear entirely.

The asymmetry in product quality that is a feature of the second model implies that the number of product types exceeds effective product variety. In particular, it has been shown that when stronger spillovers induce an increase in the heterogeneity of product sales, this will negatively affect the measure of effective product variety that was discussed in chapter 5.

Concluding, the models of economic growth introduced in this chapter provide some new perspectives on economic growth and demonstrate that building complex structures of knowledge spillovers – a strategy pursued in the literature surveyed in chapter 4 – is not the only way by which models of economic growth can be brought closer to the data. The next chapter will focus on another determinant of product variety: economic integration.

## 7.A Endogenous skill growth

The assumption of section 7.2 that the average level of skills grows at an exogenous and constant rate,  $\hat{s} = g_s > 0$ , has been made for analytical convenience. However, in real life education is not free and therefore growth in skills requires growth in resources devoted to education. This appendix discusses two cases for which constant growth in average skills is feasible in the steady state.

In order to avoid notational changes in the model, assume that the population,  $P$ , consists of the normal workforce,  $L$ , and the part of the population being a teacher or student,  $P_s$  ( $P = P_s + L$ ). Furthermore, suppose that the change in average skills is affected by the amount of human capital per capita that is available for education,  $H_s/P$ , and by a discount factor,  $\delta$ . In particular, the change in average skills is given by  $\dot{s} = H_s/P - \delta s$ . Human capital depends on the people involved in education activities and on their average education, which is assumed to equal that of the population:  $H_s = s^\varepsilon P_s$ . (Better educated teachers will teach more effectively; better educated students will learn quicker.) Substituting for  $H_s$  and dividing by  $s$  gives an expression for  $\hat{s}$ .

$$\hat{s} = s^{\varepsilon-1} \frac{P_s}{P} - \delta \quad (7.68)$$

Define  $\Lambda_s \equiv P_s/P$  and take the growth rate of  $(\hat{s} + \delta)$  to get

$$d \ln (\hat{s} + \delta) / dt = (\varepsilon - 1) \hat{s} + \hat{\Lambda}_s. \quad (7.69)$$

This last expression implies that there can be two specifications that allow for a constant and positive growth rate of skills in the steady state. First,  $\varepsilon = 1$  in combination with  $\hat{P}_s = \hat{P}$  yields  $g_s = \Lambda_s^* - \delta$ . With this specification, skill growth stems entirely from the positive effect of skills as an input on skills as an output, while the proportion of people involved in education remains constant. This specification has been proposed by Lucas (1988) and Rosen (1976). See also Arnold (1998).

Second, if  $\varepsilon < 1$ ,  $d \ln (\hat{s} + \delta) / dt$  will go to zero as time proceeds. Setting  $d \ln (\hat{s} + \delta) / dt = 0$  yields  $\hat{s} = \frac{1}{1-\varepsilon} \hat{\Lambda}_s$ . Steady state skill growth can only be positive if the proportion of the population active in education is growing, but for this proportion to grow at a constant rate, the population should grow at a different rate than the workforce. If both  $\hat{\Lambda}_s$  and  $\hat{L}$  are to be constant, the population should grow according to  $\hat{P} = g_{\Lambda_s} \frac{P_s}{L} + g_L$ .

The results presented above demonstrate that the growth rate of skills can be positive and constant in the steady state, but only under very restrictive assumptions. A more detailed and general treatment of the effects of schooling on economic growth is given by Bils and Klenow (2000).



## 7.B Endogenous growth with imperfect improvements

Suppose the quality of a product is given by  $e^{\nu i}$  with  $\nu > 0$  in stead of  $i^\nu$ , then equation 7.48 transforms into

$$x(i) = \frac{e^{\frac{\nu}{1-\gamma}i}}{\int_0^n e^{\frac{\nu}{1-\gamma}i'} di'} bL_x. \quad (7.70)$$

Substitute this into the (modified) consumption index to get

$$C = \left( \int e^{\frac{\nu}{1-\gamma}i} di \right)^{\frac{1-\gamma}{\gamma}} bL_x = \left( \frac{1-\gamma}{\nu} \right)^{\frac{1-\gamma}{\gamma}} \left( e^{\frac{\nu}{1-\gamma}n} - 1 \right)^{\frac{1-\gamma}{\gamma}} bL_x \quad (7.71)$$

The growth rate of consumption depends on the growth of product variety, and the growth of the workforce.

$$\hat{C} = \frac{\nu \dot{n}}{\gamma \left( 1 - e^{\frac{\nu}{1-\gamma}n} \right)} + \hat{L}_x \quad (7.72)$$

In the absence of population growth, steady state growth requires that  $\dot{n} = l_n$  is constant and  $\hat{L}_x = 0$ .

$$\hat{C}^* = \frac{\nu l_n^*}{\gamma \left( 1 - e^{\frac{\nu}{1-\gamma}n} \right)} \quad (7.73)$$

The steady state growth rate of consumption is not constant, but converges to  $\frac{\nu}{\gamma} l_n^*$  as  $n$  approaches infinity.

A growing population leads to explosive growth, just like with ordinary models of endogenous growth.

$$\lim_{n \rightarrow \infty} \hat{C} = \lim_{n \rightarrow \infty} \frac{\nu g_L n}{\gamma} + g_L = \infty \quad (7.74)$$

## 7.C Derivation of $V$

The solution for  $V$  as a function of  $n$  can be found by integrating equation 7.64. Rewrite this equation such that

$$\ln V(N) = F(0) - F(n), \quad F(i) \equiv \int \Phi i^k \ln [\Phi i^k] di, \quad k \equiv \frac{\nu}{1-\gamma} \quad (7.75)$$

The solution for  $F(i)$  is easy to find (using the software package *Mathematica*<sup>®</sup>).

$$F(i) = \frac{\Phi}{(1+k)^2} i^{1+k} \left( -k + (1+k) \ln [\Phi i^k] \right) + F_0 \quad (7.76)$$

Here,  $F_0$  is the constant of integration. Obtaining  $F(n)$  is now trivial, but obtaining  $F(0)$  is not because of the logarithm in the expression above. Letting  $i$  approach zero from above yields the following limit.

$$\lim_{i \downarrow 0} F(i) = \frac{\Phi}{1+k} \lim_{i \downarrow 0} i^{1+k} \ln [i^k] + F_0 \quad (7.77)$$

By L'Hôpital's rule we find that  $F(0) = F_0$  as long as  $k > -1$ .

$$\lim_{i \downarrow 0} \frac{\ln [i^k]}{i^{-1-k}} = \lim_{i \downarrow 0} \frac{d \ln [i^k] / di}{d i^{-1-k} / di} = \lim_{i \downarrow 0} \frac{-k}{1+k} i^{1+k} = 0 \quad (7.78)$$

The condition  $k < -1$  implies that  $\nu > \gamma - 1$  is required for a solution to be feasible, meaning that negative spillovers can not be too strong. With  $F(0)$  equaling  $F_0$ , we have  $V(N) = -F(n) = V(n)$ . Substitute for  $k$  to get equation 7.66.

## Notes

1. Appendix 7.A treats endogenous education.
2. See Arnold (1998) for a model that combines advances in education with knowledge spillovers.
3. The way in which the distribution of skills is modelled resembles the specification of the distribution of jobs in Dupuy and Marey (2004).
4. The consumption function corresponds to equation 4.8 with  $\delta = 1$  and  $\nu = 0$ .
5. The model can be extended with other types of jobs. The inclusion of, for example, managerial or professional jobs avoids the implication of the model that all the smart people end up doing research. The analysis of chapter 4 suggests that this modification will not lead to qualitatively different outcomes.
6. See section 4.1 for a discussion of steady states in semi-endogenous growth models.
7. For economic growth to be truly neo-endogenous, the growth rate of education should itself be endogenous, it should be positive in the absence of population growth, and it should be non-explosive when population growth is positive.
8. The precise condition for  $d \ln \Lambda^* / d\sigma < 0$  is

$$\left(1 - \frac{\alpha(\sigma\beta + 1)}{\beta(\sigma\alpha + 1)}\right) \frac{\Theta^2}{1 + \Theta} < -\ln \left(\frac{\Theta}{1 + \Theta}\right).$$

This follows from differentiating 7.28.

9. As equation 7.42 is similar to equation 7.40, it is possible to find the condition under which  $\Lambda^\# < \Lambda^{*\#}$ . This condition turns out to be  $\sigma < \frac{\beta - \alpha}{\alpha(1 - \beta)}$ .
10. This and other results that are discussed in this section can be derived following the same procedures as in chapter 4.
11. Even though profits of individual firms might decline over time, the Ramsey rule still applies as a non-autonomous change in profits is accounted for in equation 4.49. The predatory, or business-stealing, effect of the introduction of a new product is assumed to be exogenous to the individual household.



## Economic integration

Section 3.1 illustrated that economic integration is a pervasive phenomenon in the OECD-countries. The current chapter analyzes the theoretical implications of this trend for welfare. In particular, a model of an economy is presented where products are invented and patented, and where production involves fixed costs at the location of the plant. The model is used to assess the effects of instantaneous integration of a small, autarkic country into a larger economy on consumer welfare and the distribution of income. Consumer welfare in the small country rises immediately because of newly available products. Additionally, the welfare of all consumers rises due to economies of scale at the firm level. These latter benefits are gradually replaced by benefits stemming from newly invented products. The distribution of income changes due to the asymmetric distribution of patent ownership and changes in the ratio's of skilled to unskilled workers. The theoretical results presented here are complementary to those of chapters 4 and 7, which discussed how population growth, knowledge spillovers, and education can affect product variety. The model presented below will prove to be helpful for understanding the empirical evidence of chapter 10.

### 8.1 Introduction

One of the trends that characterized the period after World War II, and in particular the 1990s, has been the widespread adoption of more liberal trade and investment regimes by economically advanced countries. The most direct consequence of the more liberal economic policies is that it enables firms to operate in markets that were previously inaccessible to them. For individual firms, a more liberal economic climate brings about both opportunities and threats. Being able to sell a product in a larger number of countries provides firms with an opportunity to spread their fixed costs over a larger quantity of products, hence raising their efficiency. Liberalization also poses a threat to firms because foreign competitors may enter their domestic market.

When a firm produces a product to which there exists no close substitute, the opportunities provided by liberalization outweigh the threats. Firms that have a (natural) monopoly over a product gain access to new customers while not experiencing fiercer competition. A larger market allows these firms to increase their revenue and – in the presence of scale economies – to enhance their profitability.

The major benefit of liberalization, however, is not a better exploitation of scale economies at the firm level. Far more important is the fact that larger markets potentially provide a greater incentive for the invention and development of new products. This greater incentive for research hinges on the extent to which inventions are protected from imitation. Access to a market in which no profits can be made because of imitation by competitors is worthless to a firm and does not contribute to a greater incentive for research.

For economic liberalization to stimulate technological change it is important that the intellectual property rights (IPRs) of any firm are protected in as many countries as possible. In order to achieve this, the members of the WTO reached an agreement on the trade related aspects of intellectual property rights (TRIPS). The objective of TRIPS is stated in article 7 of the agreement:

The protection and enforcement of intellectual property rights should contribute to the promotion of technological innovation and to the transfer and dissemination of technology, to the mutual advantage of producers and users of technological knowledge and in a manner conducive to social and economic welfare, and to a balance of rights and obligations. (World Trade Organization 1994, Annex 1C)

The principle way by which this objective is to be pursued is stated in article 3, paragraph 1 of the treaty.

Each Member shall accord to the nationals of other Members treatment no less favourable than that it accords to its own nationals with regard to the protection of intellectual property (...) (World Trade Organization 1994, Annex 1C)

Given the controversies about the protection of foreign IPRs in developing countries, the suggested relation between the enforcement of IPRs and the level of welfare deserves a closer look. This chapter assesses the effects of economic integration on welfare when IPRs are strictly enforced. It shows that aggregate welfare increases in response to economic integration but also that it may severely affect the short and medium term distribution of income within and between countries.

Protection of intellectual property ensures economic growth by the standard argument that it provides firms with an incentive to develop new products. Economic integration temporarily strengthens this incentive as producers suddenly see the market for their products increase. Technological change and thus economic growth are stimulated by economic integration in much the same way as they are stimulated by population growth.

Economic integration may have a substantial impact on the distribution of income when IPRs are enforced irrespective of the country of invention as envisaged in paragraph 3.1 of TRIPS. The principle argument can be easily explained. Suppose there exist two autarkic countries of a different size. Within each country each product is patented and is manufactured by just one firm. Then it is not unreasonable to assume that the number of product types in the large country is larger than that in the small country, and that many of the types in the small country are similar to some type in the large country.

When the two countries decide to form an economic union, some comparable patents will have been registered in both countries. These patents will only be valid for the country in which they were originally registered; the other patents will become valid for the entire economic union. Clearly, the value of the patents with limited validity will be lower than the value of the other patents.

Because the majority of firms with a fully valid patent will be located in the large country, the average value of firms in the large country will rise more than the average value of firms in the small country. Many firms in the large country will see the market for their products expand, while this is unlikely for small country firms.

The asymmetric effect of economic integration on the value of firms implies that the benefits of integration will be larger for shareholders in the large country than they will be for shareholders in the small country. The shareholders in the large country get a 'free lunch' in the form of market expansion, which is not enjoyed by shareholders in the other country. In this way, the international protection of IPRs will contribute to the proliferation of income inequality without stimulating economic growth.

Although economic integration is not particularly attractive for shareholders in the small country, this does not mean that the inhabitants of this country have nothing to gain from opening up their economy. Consumers in the small country experience a large increase in their welfare immediately after integration because they gain access to products that were previously only available in the large country. This positive effect of economic integration has been emphasized by Romer (1994) and Rutherford and Tarr (2002).

In this chapter, a model of semi-endogenous growth is presented that illustrates this and other consequences of economic integration. In the model, production involves fixed costs at the location of the plant. The economies of scale that arise from these fixed cost provide an incentive for international trade à la Krugman (1979).

Another feature of the model is that new product types have to be invented before they can be produced. This feature is not uncommon to models of international trade, but is usually confined to quality improvements. An early trade model that does incorporate variety-expanding R&D is due to Grossman and Helpman (1989). A recent model by Peretto (2003a) features R&D driven quality improvements, while product variety is determined by the size of the market relative to fixed costs – the standard Dixit and Stiglitz (1977) result.

In the current model economic integration does not cause an immediate rise in the number of firms because products first have to be invented. Rather, the initial advantage of economic integration lies in the increased scale of production. Over time, the number of products gradually increases as the enhanced profitability of firms induces research. In a nutshell, the model presented here unifies the approaches to product variety by Dixit and Stiglitz (1977) and Judd (1985b).

As economic integration effectively raises the size of the workforce, its consequences for welfare are similar to those of population growth, which links the current chapter to chapter 4. As economic integration can also bring about a change in the ratio of skilled to unskilled workers, the mechanisms described by the first model of chapter 7 also play a role here. The perspective of the current chapter differs from that of chapters 4 and 7 in that it studies the consequences of a shock in the size and composition of the workforce rather than those of gradual changes.

Two sacrifices have been made to keep the model tractable. First, trade is assumed to be costless. The introduction of trading costs appears to be difficult when profits are positive.<sup>1</sup> Second, intertemporal knowledge spillovers are not easily combined with fixed costs and are therefore omitted. Consequentially, the model is restricted to semi-endogenous growth (i.e. economic growth that ultimately depends on population growth). Factor immobility is preserved throughout this chapter.

Each product is manufactured with skilled and unskilled labor. As all products are symmetric this brings up the question of how factor price equalization (FPE) is achieved. The route to FPE that is followed here is that of vertical specialization by multinational enterprises (MNEs) in the spirit of Helpman and Krugman (1985, chaps. 12 and 13). If factor proportions differ across countries, a firm will have an incentive to split up its production process into a part that uses skilled labor and a part that uses unskilled labor. A consequence of this strategy is the existence of and trade in intermediate products. Whether these intermediate trade flows are intrafirm or whether they take place between two firms (one of which a subcontractor) does not matter here. So, instead of moving factors of production to the place where they are relatively scarce, production processes are located in such a way that the costs of production are minimal.

The distinction between skilled and unskilled workers has been made not only because economic integration affects relative wages, but also because it influences the number of products per capita. High skilled workers are allocated between production and research, which implies that the cost of inventing a new product go down if skilled labor is abundant.<sup>2</sup> Economic integration not only affects the number of product types that are available world-wide directly by enlarging the scale of the economy, but also by changing the relative abundance of skilled workers.

There is a substantial literature on international knowledge spillovers that is closer to the theory of economic growth than to the theory of international trade. A seminal contribution is due to Rivera Batiz and Romer (1991), who presented a model with an expanding variety of products. Taylor (1994) discussed the 'once-off' and continuing effects of integration on economic growth. These contributions are only remotely connected to the model presented here as the productivity of researchers is not affected by previous inventions. Instead, the (pre-integration) international transfer of knowledge is closer to that featured in a model by Grossman and Helpman (1991b), where economically less developed countries imitate products from advanced countries.

Tang and Wälde (2001) analyzed the effects of integration on the structure of competition. In their model, firms producing similar goods end up in duopolies, whereas in the model introduced below they merge, such that monopolistic competition is restored instantly.

The chapter is structured as follows. Section 8.2 presents a model of an (integrated) economy with patented products and fixed production costs. The effects on aggregate welfare of two autarkic countries of unequal size forming an economic union are discussed in section 8.3. Section 8.4 treats the implications for the distribution of income. Concluding remarks can be found in section 8.5.

## 8.2 The model

The economy is inhabited by two dynastic households.<sup>3</sup> One household comprises all unskilled workers,  $L$ , and one household has all skilled workers,  $H$ , as its members. The most important difference between skilled and unskilled labor is that the former has an alternative use. Skilled labor is partly employed for production,  $H_y$ , and partly for research,  $H_R$ . 'Lifetime' utility of the households is the discounted sum of 'instantaneous' utility  $u$  with subjective discount factor  $\rho$ .

$$U_\Lambda = \int_{\tau=t}^{\infty} \Lambda_\tau u_{\Lambda,\tau} \exp[\rho(t-\tau)] d\tau, \quad \forall \Lambda \in \{L, H\} \quad (8.1)$$

Each household derives instantaneous utility from a consumption index  $C$  that depends on the quantities consumed of  $n$  different product types.

$$u_\Lambda = \ln \left( \frac{C_\Lambda}{\Lambda} \right), \quad \forall \Lambda \in \{L, H\} \quad (8.2)$$

$$C_L + C_H = C = \left[ \int_0^n y(i)^\gamma di \right]^{\frac{1}{\gamma}} \quad (8.3)$$

Firms and product types are indexed by  $i \in [0, n]$ . Products are produced in two stages. The first stage involves only unskilled labor,  $l$ , and returns intermediate products. At the second stage, skilled labor,  $h$ , and stage 1 products,  $\tilde{y}$ , are used to produce stage 2 products,  $y$ .<sup>4</sup> Products require fixed expenditure at both stages:

$$\tilde{y}(i) = \tilde{a}l(i) - \tilde{b} \quad (8.4)$$

$$\begin{aligned} y(i) &= \tilde{y}(i) \left( \frac{ah(i) - b}{\tilde{y}(i)} \right)^\beta = \tilde{y}(i)^{1-\beta} (ah(i) - b)^\beta \\ &= (\tilde{a}l(i) - \tilde{b})^{1-\beta} (ah(i) - b)^\beta \end{aligned} \quad (8.5)$$

$\tilde{a}$  and  $a$  are parameters;  $\tilde{b}$  and  $b$  are the fixed costs at the first and second stage of production, respectively.<sup>5</sup>

Each product type is patented and is manufactured by only one firm, but the two stages of production do not necessarily take place in the same country. Firms can be ‘binational’. The underlying idea is that a firm buys a patent on a product that is designed keeping in mind that it can be costly to gather all the inputs in a single place. More specifically, a part of the product is assumed to be producible separately and to require only unskilled labor. This intermediate product can be shipped without costs to any country within the economic region where it is ‘upgraded’ to a consumable good with the use of skilled labor.

Profit maximization leads to prices being a mark-up over marginal costs.

$$p_y = \gamma^{-1} \frac{w \left( l_y - \frac{\tilde{b}}{\tilde{a}} \right) + \omega \left( h_y - \frac{b}{a} \right)}{y} \quad (8.6)$$

The wage rate of unskilled labor is denoted by  $w$ , the wage rate of skilled labor is  $\omega$ . Because firms are symmetric the prices of all products are equal. The ratio of the wage rates is

$$\frac{w}{\omega} = \frac{1-\beta}{\beta} \frac{h_y - \frac{b}{a}}{l_y - \frac{\tilde{b}}{\tilde{a}}}. \quad (8.7)$$

The consumption index is taken to be the numeraire, implying that total expenditure is equal to  $C$ .

$$C = n^{\frac{1}{\gamma}} y = np_y y \Rightarrow p_y = n^{\frac{1-\gamma}{\gamma}} \quad (8.8)$$

Although neither trade nor migration exists between autarkic economies, they are assumed not to function in complete isolation from the rest of the world. In particular, a certain degree of knowledge spillovers occurs between economies. Researchers will imitate a product that has been invented in another economy rather than try to invent a product that is entirely new – simply because imitation is cheaper. The flow of patents, each representing a design of a new product type,



depends on the number of researchers employed and on whether the product is an imitation or a true invention.<sup>6</sup> A researcher can produce one design at the time or he can produce  $\delta > 1$  imitations. Imitation is only possible when the number of product types in the world,  $n_w$ , is strictly larger than the number of types available 'domestically'. Summarizing what has been stated above, the flow of new products is given by:

$$\dot{n} = \begin{cases} H_R & \text{if } n = n_w \\ \delta H_R & \text{if } n < n_w \end{cases} \quad (8.9)$$

The dot denotes the derivative with respect to time.

Entry and exit are free in the research sector thus profits in the research sector are zero. The value of a design,  $v$ , is limited by the wage a researcher receives.

$$v \leq \omega \quad (8.10)$$

The distribution of total revenues from production can be found using the wage ratio and both expressions for the price (see section 4.5).

$$C = wL + \omega H_y + n\pi \quad (8.11)$$

$$wL = (1 - \beta) \gamma C + wn \frac{\tilde{b}}{a} \quad (8.12)$$

$$\omega H_y = \beta \gamma C + \omega n \frac{b}{a} \quad (8.13)$$

$$n\pi = (1 - \gamma) C - n \left( w \frac{\tilde{b}}{a} + \omega \frac{b}{a} \right) \quad (8.14)$$

Using these last results, the profit to skilled wage ratio can be expressed as

$$\frac{\pi}{\omega} = (1 - \gamma) \frac{C}{n\omega} - \frac{w}{\omega} \frac{\tilde{b}}{a} - \frac{b}{a} \quad (8.15)$$

$$= \frac{(1 - \gamma) (H_y - n \frac{b}{a})}{\beta \gamma n} - \frac{(1 - \beta) \tilde{b} H_y - n \frac{b}{a}}{\beta \tilde{a} L - n \frac{\tilde{b}}{a}} - \frac{b}{a} \quad (8.16)$$

where  $\pi$  represents the profits of a firm.

In accordance with the results established in section 4.4, the Ramsey rule is the same for both households.

$$\hat{C}_L - \hat{L} = \hat{C}_H - \hat{H} = \hat{C} - g_P = \frac{\pi}{v} - \tilde{\rho}. \quad (8.17)$$

Variables wearing a hat are growth rates; the (constant) growth rate of the work-force is  $g_P$ . The total population,  $P$ , is the sum of  $L$  and  $H$ . Substitute for  $\hat{C}$  and  $\pi/v$  in the Ramsey rule to get the optimal growth rate of consumption, provided that research is feasible ( $v = \omega$ ).

$$\hat{C} = \frac{(1 - \gamma) (H_y - n \frac{b}{a})}{\beta \gamma n} - \frac{(1 - \beta) \tilde{b} H_y - n \frac{b}{a}}{\beta \tilde{a} L - n \frac{\tilde{b}}{a}} - \frac{b}{a} - \tilde{\rho} + g_P \quad (8.18)$$

The total consumption can be written in terms of  $n$ ,  $L$ , and  $H_y$  using equations 8.3 and 8.5.

$$C = n^{\frac{1}{1-\gamma}} y = n^{\frac{1-\gamma}{\gamma}} \left( \tilde{a} L - n \tilde{b} \right)^{1-\beta} (a H_y - n b)^{\beta} \quad (8.19)$$

The balanced growth rate can be found in a straightforward manner. The equations for consumption (8.19) and consumption growth (8.18) show that the growth rate of consumption can only be constant if  $\hat{L} = \hat{H} = g_P$ . Equations 8.9 and 8.19 reveal that the balanced growth rate of consumption is  $\frac{1}{\gamma}g_P$ . It immediately follows that per capita consumption growth is  $\frac{1-\gamma}{\gamma}g_P$ , even though knowledge spillovers are absent. The model is characterized by 'semi-endogenous' growth, that is, economic growth depends crucially on population growth. That the model is saddle path stable is shown in the appendix.

### 8.3 Aggregate welfare

The approach to economic integration adopted in this chapter is that of a 'big bang' unification of two economies: a large one, labeled *A*, and a small one, labeled *B*. Before unification, both economies are autarkic. After unification, the two economies form a single integrated economy.

According to the model of the previous section, the number of products is linearly related to the size of the population in the steady state. This implies that, before unification, the large country generally has more product types than the small country.<sup>7</sup> What happens to the number of types in each country after unification depends on how similar the two sets of types are.

The larger size of country *A* ensures that the flow of new product designs is larger than in country *B*, assuming that both countries are in their steady state. This difference in the flow of new products (eventually) causes the number of products in *A* to exceed the number of products in *B*. For this reason, all researchers in country *B* will be busy imitating designs invented in country *A*. As a result, the products that are available in country *B* are a subset of the products available in *A*. After the economies of the two countries have integrated there will no longer be imitation in any of the two countries.

Immediately after unification, the set of product types available for consumption in the newly formed economy equals the set of types that is produced in the large country. Consumers in the small country will see an immediate expansion in the number of types, whereas consumers in the large country will not see an immediate change.

There is a slight complication arising from the fact that a subset of product types is patented in both the large and the small country – presumably by different firms. The natural way of dealing with this situation is to preserve the rights attached to the patent only for the country it was valid for preceding the unification. Firms originally established in country *B* will still be allowed to sell their products domestically, but it will not be possible for them to sell in country *A*. Firms in country *A* can sell their products in both countries, provided that the product has not been patented in country *B* prior to unification.

The duopoly that arises for some products is not stable. Increasing returns to scale at the firm level provides an incentive for the shareholders of two firms owning a patent on the same product to force a merger of the two firms. Consequently, unification triggers an instantaneous wave of 'cross-border' mergers.

The occurrence of mergers raises the issue of how dividends should be distributed among shareholders from the two countries. The most neutral way of handling this problem is by assuming a 'fair deal' such that dividends are allocated according to the size of the future market of each country. In particular, if  $n_{B,0}$  is the number of firms that merge and  $\kappa_i^*$  is the long term share of country

$\iota$  in the aggregate income of the integrated economy, then the dividends received by country  $\iota$  are  $n_{B,0}\pi\kappa_\iota^*$ .

Another issue concerns the effects of unification on the location of production. If the ratio of skilled to unskilled workers is the same in both countries, no plants will be relocated; only the redundant plants of merged firms will be closed. When the ratio of skilled to unskilled workers is not the same in both countries, some firms will find it profitable to engage in vertical specialization, separating the production of intermediates from that of final goods. A part of the production of intermediates will then be shifted to the country with a relative abundance of unskilled workers, while a part of the production of final goods moves in the opposite direction.

The relocation of plants ensures that, given the type of labor, wage rates are identical across the entire integrated economy. If this would not be the case, a firm could increase its profits by relocating a part of its production activities. Hence, the integrated economy is truly integrated even though labor migration is absent. It is the movement of economic activity rather than the movement of production factors that ensures FPE.

Having settled the issues concerning property rights and FPE, we can now turn to the effects of economic integration on income and consumption. After the transition path of the integrated economy has been found (see appendix), the consumption paths of the households of the two countries have to be derived. According to the Ramsey rule (equation 8.17) consumption grows at the same pace in both countries. This implies that the allocation of consumption between the two countries is constant over time.

Ruling out unsustainable debts, each country's share in total consumption should not exceed its share in long-term income,  $C_{it}/C_t \leq \kappa_\iota^*$ . Because this condition has to hold for all countries simultaneously,  $C_{it} = \kappa_\iota^* C_t$ . Once we know  $\kappa_\iota^*$ , we can solve for the consumption paths of both countries. Aggregate income in country  $\iota$  is given by

$$Y_{it} = \omega_t H_{it} + w_t L_{it} + \pi_t \left( W_{i0} + \int_0^t \frac{Y_{t\tau} - C_{t\tau}}{Y_\tau - C_\tau} \dot{n}_\tau d\tau \right), \quad (8.20)$$

where  $W_{i0}$  is the equivalent in number of firms of shareholdings by country  $\iota$  immediately after the mergers. As  $\lim_{t \rightarrow \infty} \frac{Y_{it} - C_{it}}{Y_t - C_t} = \lim_{t \rightarrow \infty} \frac{Y_{it}}{Y_t} = \kappa_\iota^*$ ,<sup>8</sup> we have

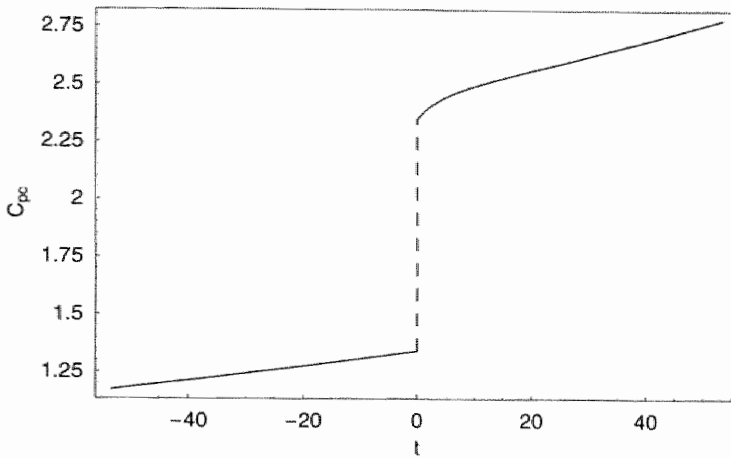
$$\kappa_\iota^* = \lim_{t \rightarrow \infty} \frac{\omega_t H_{it} + w_t L_{it}}{Y_t} + \lim_{t \rightarrow \infty} \frac{\pi_t (W_{i0} + (n_t - n_{sd}) \kappa_\iota^*)}{Y_t}. \quad (8.21)$$

In the last term,  $sd$  is the period after which distribution of income is (approximately) stable. The first term is constant as soon as the integrated economy is in the steady state. In the second term, both  $W_{i0}$  and  $n_{sd}$  are asymptotically irrelevant as they are constants. What is left over, provides us with a solution for  $\kappa_\iota^*$ :

$$\kappa_\iota^* = k_{0\iota} + \lim_{t \rightarrow \infty} \frac{\pi_t n_t \kappa_\iota^*}{Y_t} = k_{0\iota} + k_1 \kappa_\iota^* = \frac{k_{0\iota}}{1 - k_1} \quad (8.22)$$

The constants  $k_{0\iota}$  and  $k_1$  follow from the steady state solution for the integrated economy.

Figure 8.1 displays consumption per capita of the small country before,  $t < 0$ , and after,  $t > 0$ , the unification. Clearly, unification has an immediate and large impact on the consumption of the small country's inhabitants. Consumption rises for two reasons. First, unification yields immediate access to the product types



**Figure 8.1.** Per capita consumption – small country

that previously were only available in the large country. This effect is clearly visible in figure 8.2 and has been discussed by Romer (1994) and Rutherford and Tarr (2002). Second, the larger markets for products brought about by unification enables further exploitation of economies of scale due to fixed costs. Figure 8.3 displays the scale of production for final goods. For both countries the scale of production rises instantly after the unification, but it declines gradually as new product are being invented. Because the small country no longer imitates after integration, the steady state number of produced types per capita declines in this country. This is why the steady state scale of production in the small country increases after integration. The steady state scale of production of the large country is not affected by integration.

Unlike the inhabitants of the small country, the inhabitants of the large country do not seem to have any immediate benefits from unification (figure 8.4). For the large country, unification does not bring about a spectacular expansion of the set of product types, although the growth of the number of types does temporarily accelerate as can be seen from figure 8.5.

The increase in scale caused by unification has a positive effect on profits. Higher profits, in turn, induce people to increase their savings rate – hence the initial decline in consumption in figure 8.4 – and causes the value of shares to inflate. High share prices make research more attractive and cause the rate of invention to accelerate.

The acceleration in the rate of introduction of new products is not sustainable. The high rate of invention has a depressing effect on the scale of production and, consequently, on profits. This causes the number of researchers relative to the size of the population to decline until the steady state is reached.

Concluding, unification leads to an immediate increase in product variety for the inhabitants of the small country while simultaneously raising the scale of production in both countries. The subsequent rise in profits stimulates research and induces the invention of new goods.

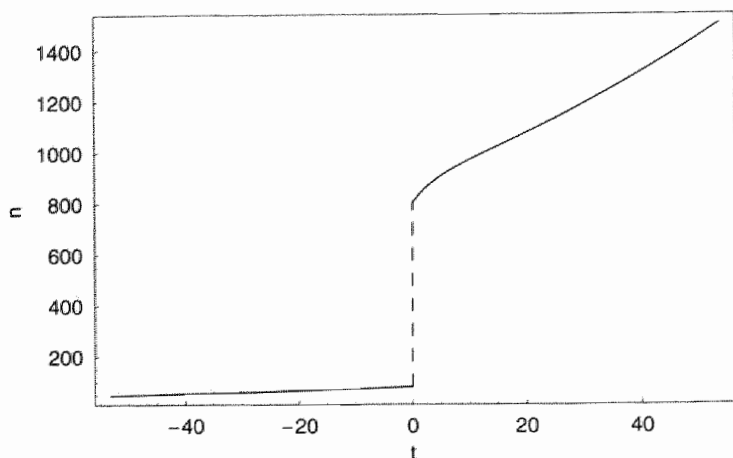


Figure 8.2. Number of types consumed – small country

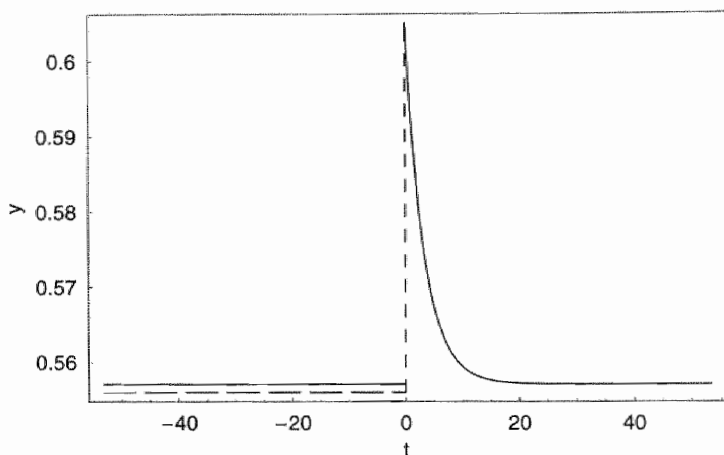
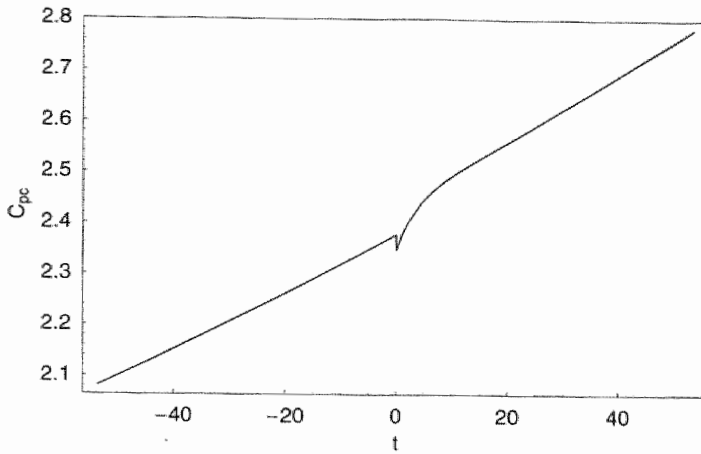


Figure 8.3. Scale of final goods production for small country (dashed) and large country (solid)

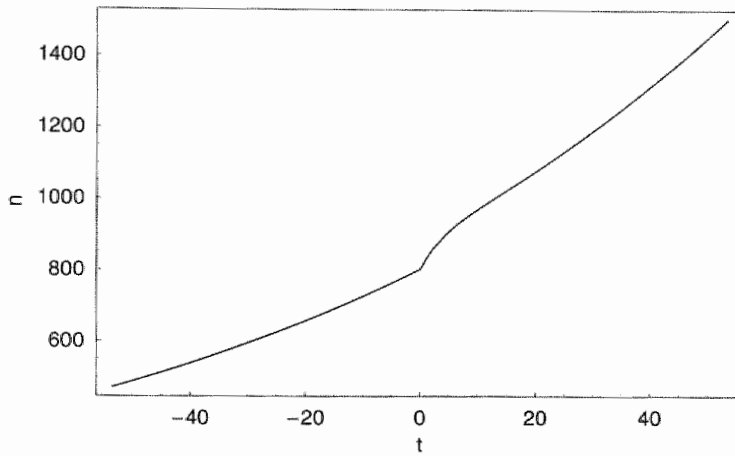
## 8.4 Income inequality

In the integrated economy, four groups of households can be distinguished: skilled workers in  $A$ , unskilled workers in  $A$ , skilled workers in  $B$ , and unskilled workers in  $B$ . Each of these groups can be affected in different ways by economic integration. This section studies the income and consumption paths of the four groups for two scenario's. The first scenario is that of the previous section where a small economy merges with a large economy that has the same ratio of skilled to unskilled workers. In the second scenario, the small economy has an abundance of unskilled labor relative to the large economy.

The consumption path for each group can be found in a manner analogous to the procedure that was used to find the consumption paths of the two countries. By the Ramsey rule the share in aggregate consumption of group  $j$  is constant over time, equalling  $\kappa_j^*$ . This property facilitates the derivation of the distribution of income.



**Figure 8.4.** Per capita consumption – large country



**Figure 8.5.** Number of types consumed – large country

Infinite horizon utility maximization causes consumption paths to be governed entirely by the long term distribution of income. This obscures some of the distributional effects of integration that would occur in real life. To get a better grasp of the short and medium term effects of unification it is useful to take a look at the income paths of the four groups.

Unfortunately, these income paths still have to be recovered. Income for the individual groups depends on wages, on initial shareholdings, and on shareholdings accumulated after  $t = 0$ .

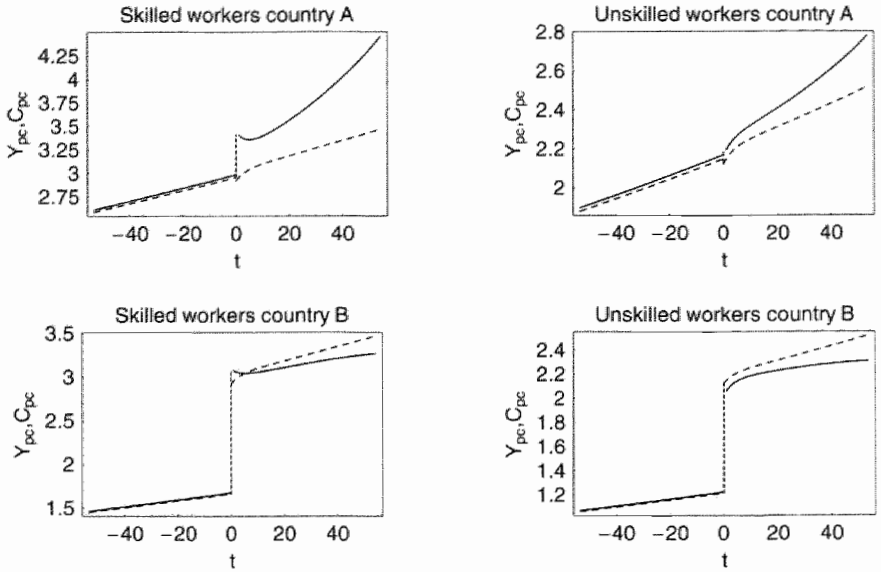


Figure 8.6.  $Y_{pc}$  (solid) and  $C_{pc}$  (dashed) by income group

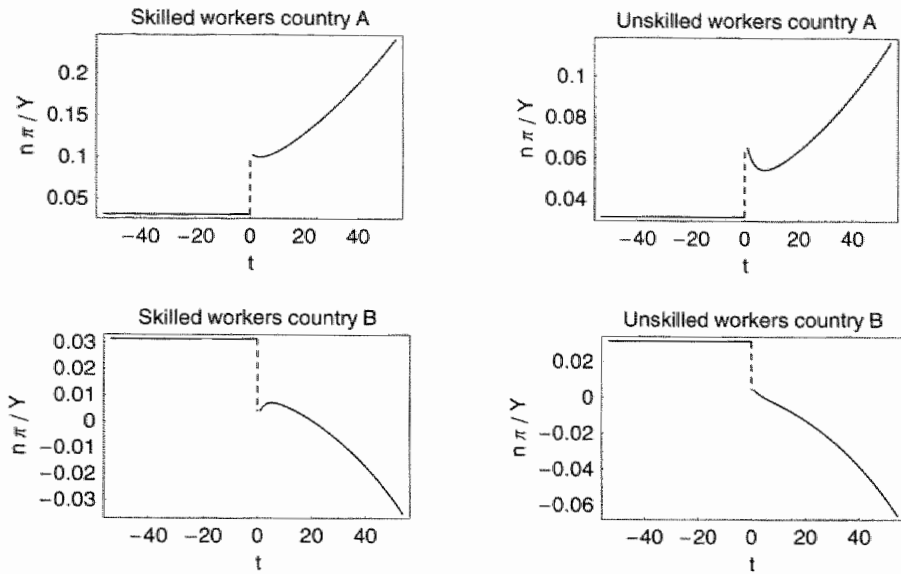
$$\begin{aligned}
 Y_{HA,t} &= \omega_t H_{A,t} \\
 &\quad + \pi_t \left( \kappa_{0HA}^* [\kappa_A^* n_{B0} + (n_{A0} - n_{B0})] + \int_0^t \frac{Y_{HA,\tau} - C_{HA,\tau}}{Y_\tau - C_\tau} \dot{n}_\tau d\tau \right) \\
 Y_{LA,t} &= w_t L_{A,t} \\
 &\quad + \pi_t \left( \kappa_{0LA}^* [\kappa_A^* n_{B0} + (n_{A0} - n_{B0})] + \int_0^t \frac{Y_{LA,\tau} - C_{LA,\tau}}{Y_\tau - C_\tau} \dot{n}_\tau d\tau \right) \\
 Y_{HB,t} &= \omega_t H_{B,t} + \pi_t \left( \kappa_{0HB}^* \kappa_B^* n_{B0} + \int_0^t \frac{Y_{HB,\tau} - C_{HB,\tau}}{Y_\tau - C_\tau} \dot{n}_\tau d\tau \right) \\
 Y_{LB,t} &= w_t L_{B,t} + \pi_t \left( \kappa_{0LB}^* \kappa_B^* n_{B0} + \int_0^t \frac{Y_{LB,\tau} - C_{LB,\tau}}{Y_\tau - C_\tau} \dot{n}_\tau d\tau \right)
 \end{aligned}$$

In the expression stated above,  $\kappa_{0j}^*$  is the long term share of group  $j$  in the aggregate income of the corresponding autarkic economy before unification. Patents ranging from  $n_{B0}$  to  $n_{A0}$  are only possessed by firms in country A.

As the level of income at  $t > 0$  depends on the income of previous periods, income paths can only be derived recursively. For the simulation results presented below,  $\int_0^t \frac{Y_{j\tau} - C_{j\tau}}{Y_\tau - C_\tau} \dot{n}_\tau d\tau$  has been replaced by  $\sum_{\tau=1}^t \frac{Y_{j\tau-1} - C_{j\tau-1}}{Y_{\tau-1} - C_{\tau-1}} \Delta n_\tau$ .

Figure 8.6 illustrates the effects of unification on per capita income and consumption.<sup>9</sup> The ratio of skilled to unskilled workers is the same in both countries.

The figure clearly visualizes three phenomena. Without a doubt, the most important phenomenon is the increase in welfare for the inhabitants of country B caused by access to a wider set of products from  $t = 0$  onwards. This has already been discussed in the previous section.



**Figure 8.7.** Net asset income

A second phenomenon is the jump in the income of skilled workers in country *A*. This group owns the majority of the shares in country *A* firms. After unification, most of these firms can suddenly sell to consumers in *B* because they own patents on products that were not registered in country *B*. The corresponding rise in profits is mainly appropriated by skilled workers in *A*.

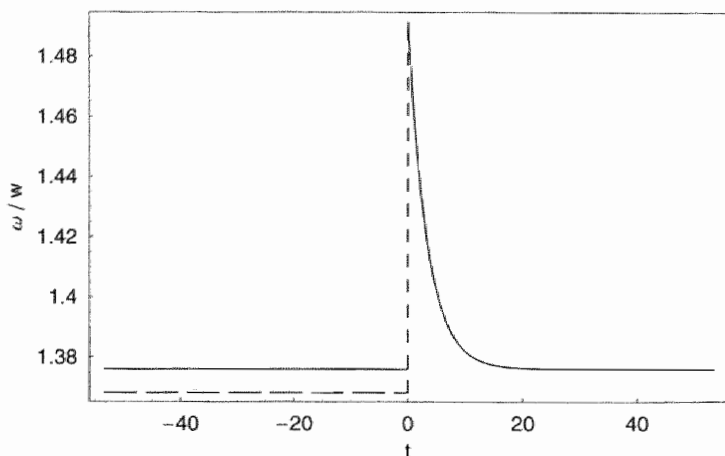
Third, consumption by households in country *B* exceeds their income. These households can lend from households in country *A* because their current share in income is below its long term value. This implies that the debt accumulated by country *B* will vanish asymptotically and is therefore sustainable.

Figure 8.7 shows net income from the possession of assets for all four groups. Especially skilled workers in *A* see their income from assets increase in response to the integration of the two economies. Unskilled workers in country *B* become net debtors very quickly (around  $t = 5$ ).

As both countries have the same ratio of skilled to unskilled workers, the effects of unification on the ratio of wages,  $\omega/w$ , are largely transitory. This can be seen in figure 8.8. The temporary rise in the wages of skilled workers relative to the wages of unskilled workers is due to the increased attractiveness of research. Once the steady state has been reached, the ratio of wages has returned to the pre-unification level, at least in the large country. The steady state wage ratio in the small country rises somewhat as products can no longer be imitated, which reduces the effective abundance of skilled workers.

When the two integrating economies have an identically composed labor force, virtually all persistent changes in the distribution of income are due to the unequal distribution of shares between small country shareholders and large country shareholders. After unification, the shareholders of the small country suddenly have a minority stake in just a handful of firms. This contrasts sharply with the pre-unification era during which domestic shareholders controlled all the firms sell-





**Figure 8.8.** Wage inequality for small country (dashed) and large country (solid)

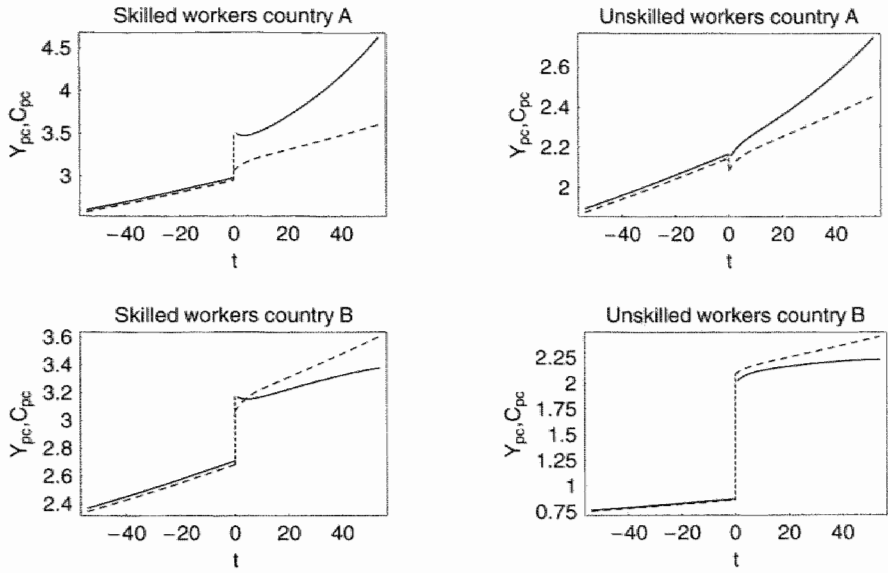
ing products on the domestic market. It also contrasts sharply with gains from unification for shareholders in the large country.

Next, we will turn to the second scenario in which the small country's ratio of skilled to unskilled workers is much lower than that of the large country. In particular, for the small country  $H_B/L_B = 0.2$  while for the large country this ratio is kept at 0.4. The other parameters are the same as with the first scenario, such that the ratio  $H/L$  for the entire post-unification economy equals 0.38.

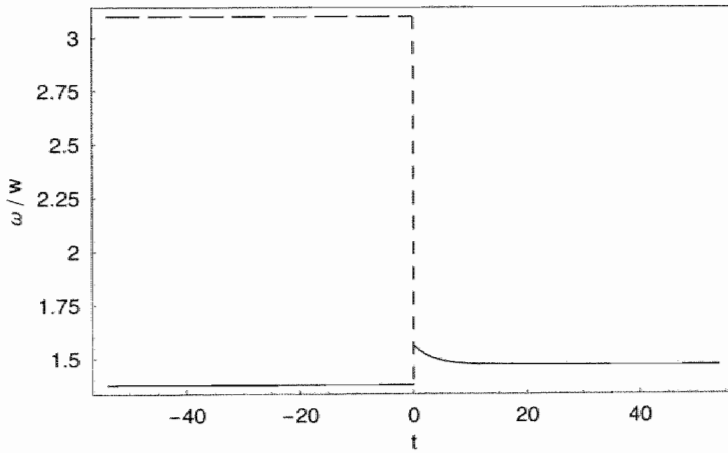
Figure 8.9 shows income and consumption per capita for the four groups. In comparison with the previous scenario, unskilled workers in  $B$  gain more from integration than before as do the skilled workers in  $A$ , whereas there do not seem to be any gains anymore for unskilled workers in  $A$ . The gains for skilled workers are less than with the first scenario.

All these deviations from the first scenario are the result of the change in the ratio of wages. Figure 8.10 shows that unification has permanent effects on the ratio  $\omega/w$  for both countries. This is not surprising as the relative ratio of skilled to unskilled workers after unification is 0.38 in stead of 0.4. In the small country  $\omega/w$  falls sharply after unification, while  $\omega/w$  rises somewhat in the large country.

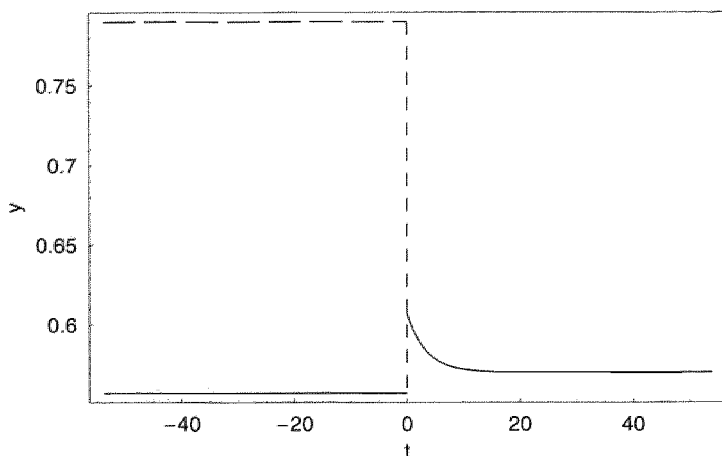
Besides the direct effect of the change in the wages, there is also an indirect effect that severely reduces the gains from integration for the shareholders in country  $B$ , mostly skilled workers. As can be seen from figures 8.3 and 8.11, the rise in  $H/L$  has affected the scale of production. The scarcity of skilled labor had worked as a 'barrier to invention' – or rather, a 'barrier to imitation' – thereby leading to fewer product types per capita and higher profits than is the case in scenario one. Due to the disappearance of this 'barrier to invention', the shareholders in country  $B$  have less to gain from unification than in the first scenario.



**Figure 8.9.**  $Y_{pc}$  (solid) and  $C_{pc}$  (dashed) by income group; shock in  $H/L$



**Figure 8.10.** Wage inequality for small country (dashed) and large country (solid); shock in  $H/L$



**Figure 8.11.** Scale of final goods production for small country (dashed) and large country (solid); shock in  $H/L$

## 8.5 Conclusion

Economic integration enables consumers in small countries to buy a larger variety of products. Additionally, the welfare of all consumers rises due to economies of scale at the firm level. The corresponding rise in profits makes research more attractive, which leads to the invention of new products. In this way, the benefits due to economies of scale are gradually replaced by benefits stemming from newly invented products. Economic integration gives therefore rise to the same scale effects as population growth does.

Although economic integration stimulates aggregate welfare, the gains from integration are not likely to be distributed evenly. In the typical situation where a small country with many unskilled workers joins a large economy, changes in the distribution of income are only partly caused by the change in relative wages.

The enforcement of IPRs has two effects on the distribution of income. First, not all shareholders benefit equally from integration due to the asymmetric validity of patents. All of the wealth of small country shareholders is invested in patents with limited geographical validity, while a part of the patents owned by large country shareholders are valid in both countries. By this mechanism, the international protection of IPRs will contribute to the proliferation of income inequality. Second, the 'barrier to invention' that is due to the scarcity of skilled workers in the small country is reduced after unification, which negatively affects the value of patents in the small country.

In reality the consequences of rapid economic integration might be understated by the model. First, patents on imitated products might be ruled illegal, thereby strengthening the negative effects of integration on the relative income of small country shareholders. Second, firms in the larger economy will tend to produce goods with a higher quality, such that consumers in the small country will not only see the range of available products expand but these products will also be of a higher quality.

Thus, economic integration has a profound influence on the number of types that a country produces as well as on the number of types that are available for consumption in that country. The results presented above are complementary to

the results of chapters 4 and 7, and will prove to be helpful for understanding the empirical evidence of chapter 10.

## 8.A Stability of the model

A first step in the search after the path leading from an initial state of the economy to balanced growth involves the rescaling of variables such that the rescaled variables are constant along the balanced growth path. Technically, rescaling transforms a moving equilibrium into a fixed equilibrium.

$$c \equiv \frac{C}{P^{\frac{1}{\gamma}}}, \quad l \equiv \frac{L}{n} - \frac{\tilde{b}}{a}, \quad \nu \equiv \frac{n}{P}, \quad h \equiv \frac{H}{n} - \frac{b}{a}, \quad h_R \equiv \frac{H_R}{n}$$

The rescaled variables  $l$  and  $h$  refer to the variable part of the employment by a single firm and should not be confused with  $l_y$  and  $h_y$  which refer to total employment by a firm.

After rescaling equations 8.9, 8.19, and 8.18 become

$$\hat{\nu} = h_R - g_P \quad (8.23)$$

$$c = \tilde{a}^{1-\beta} a^\beta \nu^{\frac{1}{\gamma}} l^{1-\beta} (h - h_R)^\beta \quad (8.24)$$

$$\hat{c} = \frac{(1-\gamma)}{\beta\gamma} (h - h_R) - \frac{(1-\beta)\tilde{b}(h - h_R)}{\beta\tilde{a}l} - \frac{b}{a} - \tilde{\rho} + \frac{\gamma-1}{\gamma} g_P. \quad (8.25)$$

Moreover,  $l$  and  $h$  can be expressed as functions of  $\nu$

$$l = \frac{1}{(1+\varphi)\nu} - \frac{\tilde{b}}{a} \quad (8.26)$$

$$h = \frac{\varphi}{(1+\varphi)\nu} - \frac{b}{a} \quad (8.27)$$

where  $\varphi \equiv H/L$ . Except for temporary shocks, the growth rates both of skilled and of unskilled labor are taken to be equal and constant. This implies that  $\varphi$  is constant as a rule and thus that it is possible to reach the steady state.

The last five equations yield a system of differential equations in  $\nu$  and  $h_R$ . The system has a fixed point which has coordinates

$$h_R^* = g_P$$

$$\nu^* = \frac{\frac{1-\gamma}{\beta\gamma} \frac{\varphi}{(1+\varphi)}}{\frac{1}{\beta\gamma} \frac{b}{a} + \tilde{\rho} + \left( \frac{(1-\gamma)(1+\beta)}{\beta\gamma} - \frac{(1-\beta)\tilde{b}}{\beta\tilde{a}} \right) g_P + \frac{(1-\beta)\tilde{b}}{\beta\tilde{a}}}$$

With  $\nu$ ,  $h$ , and  $h_R$  constant in the steady state, the profit value ratio is also constant over time, implying that  $\tilde{\rho} = \rho$ .

The stability of the fixed point can be checked by a linear approximation to the system around the steady state. Table 8.1 reports the eigenvalues and eigenvectors stemming from the linearized system for a range of plausible parameter values. As every eigenvalue is accompanied by an eigenvalue of the opposite sign, the fixed point exhibits saddle point stability. The eigenvectors indicate that the stable and unstable manifolds do not 'switch places' when parameter values change.

The non-linearized transition path has been found using backward integration (Brunner and Strulik 2002). This technique exploits the fact that the stable manifold becomes the unstable manifold when time is running backward. A numerical solution to an unstable manifold belonging to a saddle point can be found

Table 8.1. Eigenvalues and eigenvectors for linearized system

$\beta$	$\gamma$	$g_P$	$\phi$	Eigenvalue	Eigenvector
0.1	0.5	0	0.1	11.5	(0.00712, 1)
				-1.02	(-0.0798, 1)
0.1	0.5	0	10	$0.267 \times 10^3$	$(0.179 \times 10^{-3}, 1)$
				$-0.256 \times 10^3$	$(-0.186 \times 10^{-3}, 1)$
0.1	0.5	0.04	0.1	11.9	(0.0066, 1)
				-1.06	(-0.0739, 1)
0.1	0.5	0.04	10	$0.266 \times 10^3$	$(0.179 \times 10^{-3}, 1)$
				$-0.256 \times 10^3$	$(-0.187 \times 10^{-3}, 1)$
0.1	0.99	0	0.1	25.1	$(0.183 \times 10^{-3}, 1)$
				-14.7	$(-0.313 \times 10^{-3}, 1)$
0.1	0.99	0	10	$0.925 \times 10^3$	$(0.00108 \times 10^{-3}, 1)$
				$-0.915 \times 10^3$	$(-0.00109 \times 10^{-3}, 1)$
0.1	0.99	0.04	0.1	25.1	$(0.183 \times 10^{-3}, 1)$
				-14.7	$(-0.312 \times 10^{-3}, 1)$
0.1	0.99	0.04	10	$0.925 \times 10^3$	$(0.00108 \times 10^{-3}, 1)$
				$-0.915 \times 10^3$	$(-0.00109 \times 10^{-3}, 1)$
0.9	0.5	0	0.1	-1.02	(-0.0461, 1)
				2.21	(0.0212, 1)
0.9	0.5	0	10	34.7	(0.00236, 1)
				-33.5	$(-0.00244, 1)$
0.9	0.5	0.04	0.1	-1.06	(-0.0427, 1)
				2.3	(0.0196, 1)
0.9	0.5	0.04	10	34.5	(0.00237, 1)
				-33.3	$(-0.00246, 1)$
0.9	0.99	0	0.1	11.1	$(0.0863 \times 10^{-3}, 1)$
				-9.93	$(-0.0967 \times 10^{-3}, 1)$
0.9	0.99	0	10	22.9	$(0.201 \times 10^{-3}, 1)$
				-21.7	$(-0.212 \times 10^{-3}, 1)$
0.9	0.99	0.04	0.1	11.1	$(0.0864 \times 10^{-3}, 1)$
				-9.96	$(-0.0964 \times 10^{-3}, 1)$
0.9	0.99	0.04	10	22.9	$(0.201 \times 10^{-3}, 1)$
				-21.7	$(-0.212 \times 10^{-3}, 1)$

The other parameter values are:  $\rho = 0.04$ ,  $\delta = 1$ ,  $a = \bar{a} = b = \bar{b} = 1$

easily. With a starting point close to the unstable manifold, the numerical solution will quickly converge to the actual manifold. The parameter values used in the first simulation are  $\beta = 0.3$ ,  $\gamma = 0.8$ ,  $\rho = 0.03$ ,  $g_P = 0.01$ ,  $\phi = 0.4$ , and  $a = \bar{a} = b = \bar{b} = 1$  for both countries;  $P(0) = 1000$  and  $n(0) = 50$  for country A prior to integration;  $P(0) = 100$  and  $n(0) = 5$  for country B prior to integration. For the second simulation, the following alterations were made:  $\phi = 0.4$  for country A and  $\phi = 0.2$  in country B.

## Notes

1. Peretto (2003a) did find a way to combine quality R&D in the presence of trading costs.
2. The consequences of changes in the distribution of skills are studied in detail in section 7.2.
3. The absence of migration strictly implies that all members of a household are located in the same country. This shortcut leads to the same results as when households are country specific.
4. The assumption that the first stage only involves unskilled labor, while the second stage requires skilled labor is arbitrary. The behind this assumption is that raw materials are gathered by unskilled workers, while skilled workers ('artisans') are needed to transform the raw materials into useful products. A counterexample can also be given. Skilled workers produce high-tech components are assembled into a final product by unskilled laborers. Although only the first case has been incorporated in the model, both cases are perfectly plausible. The main outcomes of the model are not affected by this simplification.
5. Fixed costs are not essential to most of the results related to welfare that are discussed below, but fixed costs do help to explain how integration affects trade and the scale of production.
6. Knowledge spillovers from previous research efforts are omitted in order to keep the dynamic behavior of the model tractable.
7. If the two countries do not differ that much in size while imitation is significantly cheaper than invention, then the number of products might be identical in both countries.
8. The share of country  $i$  in aggregate savings cannot grow indefinitely as this would imply unsustainable borrowing.
9. The parameter settings are the same as in the previous section.



## Idiosyncratic firm performance

This chapter discusses the effects of idiosyncratic shocks in the performance of firms on product variety and economic growth. Idiosyncratic and permanent productivity shocks make consumers shift their expenditure towards firms that have a high productivity and therefore low prices. This process of variation and selection leads to both growth in aggregate output and continuing loss in product variety. The negative correlation between product variety and aggregate output found in this chapter contrasts starkly with the findings of earlier chapters. Whether the overall effect of idiosyncratic shocks on consumer welfare is positive or negative, depends on the substitutability of products. An empirical test of the predicted negative effect of idiosyncrasy on product variety can be found in chapter 10.

### 9.1 Introduction

The models of economic growth presented in chapters 4, 7, and 8 all have the property that the growth of an economy is always the result of deliberate investments made by economic agents like households or firms. The reason why these agents invest is because they want to improve or maintain their standard of living or profitability in the future. According to the perspective adopted in these models, economic growth is therefore inextricably linked to forward-looking behavior. The approach taken in this chapter does not require forward-looking agents. In stead, economic growth is the result of a mechanism of variation and selection. Random events cause firms to differ with respect to their productivity, while static optimization by households ensures that efficient firms produce more than inefficient firms. This process of variation and selection raises the aggregate productivity of the economy in a way that resembles Fisher's fundamental theorem of natural selection:<sup>1</sup>

The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time. (Fisher 1958 [1930], p. 37)

Two models are presented in this chapter, a model in discrete time and a model in continuous time. Both models are characterized by monopolistic competition and by idiosyncratic and permanent shocks to the productivity of firms. A consequence of these productivity shocks is that the differences between firms will widen over time. In the context of monopolistic competition this does not necessarily



result in the exit of less efficient firms, but it implies that some firms will produce more and have higher profits than other firms. For monopolistic firms, the process of selection is gradual and continuous rather than sudden and discrete.

Variation caused by idiosyncratic shocks and the subsequent selection by market forces affect consumer welfare in two ways. On the one hand, the ongoing shift to more productive firms is beneficial to consumers in the sense that they are able to consume a larger quantity than before. On the other hand, the concentration of production with efficient firms implies a reduction in the effective variety of consumer goods, which has a depressing effect on consumer welfare. The greater the difference between products, or, the poorer their substitutability, the more consumers will suffer from this loss in product variety. In particular, the loss in variety is exactly off-set by the gain in quantity if the elasticity equals two. Below this number, welfare is adversely affected by idiosyncratic shocks, above this number the overall effect on welfare is positive. The purpose of this chapter is to show what consequences increasing firm heterogeneity has for the growth of welfare and product variety. It is not the aim of this chapter to analyze what events and mechanisms cause firms to be different.<sup>2</sup>

Nelson and Winter (1982, chap. 10), Conlisk (1989), and Metcalfe (1998, 1999) all have presented models of economic growth in which Fisher's theorem plays a prominent role. The models presented in this chapter differ from their models in that they are based on the framework of monopolistic competition that is used in standard (mainstream) growth theory. Furthermore, not only the upside of Fisher's theorem, namely productivity growth, is treated here but also its downside, the loss of variety. Both matter for consumer welfare.

Except for the absence of R&D in the discrete model, the models of this chapter are not substantially different from the models of standard growth theory. In particular, continuously increasing firm heterogeneity has been a common property of growth models ever since Grossman and Helpman's quality ladder model (Grossman and Helpman 1991c, p. 49). In fact, mechanisms similar to Fisher's theorem can be found in many models with stochastic quality improvements in more than one type of product.<sup>3</sup> This chapter highlights the working of Fisher's theorem in this class of models and shows that it has ambiguous effects on the growth of consumer welfare. Although this chapter focusses on heterogeneity in the costs of production, the conclusions that are drawn here also apply to heterogeneity in the quality of products.

The chapter is structured as follows. Section 9.2 outlines a model in discrete time and presents results for a continuum of firms. Results for a discrete number of firms are obtained via Monte-Carlo simulations and can be found in section 9.3. A model in continuous time and with an expanding range of product types is presented in section 9.4. The purpose of this model is to show how the discrete time model can be translated into a standard growth model. Section 9.5 summarizes the findings.

## 9.2 Discrete model and asymptotic results

Consider an economy where  $n$  different types of goods are produced according to  $y(i) = b(i)l(i)$ , where  $y(i)$  is the quantity of good  $i$ ,  $l(i)$  is the labor used and  $b(i)$  is a cost parameter. Each product is produced by a single firm. Consumer  $j$ 's preferences are reflected by the consumption function  $c(j) = (\sum_i y(ij)^\gamma)^{\frac{1}{1-\gamma}}$  with  $\gamma < 1$  (the elasticity of substitution,  $\varepsilon$ , equals  $1/(1-\gamma)$ ). Here,  $y(ij)$  is the quantity of good  $i$  consumed by consumer  $j$ . Consumers maximize their consumption

subject to a budget constraint. Assuming that all consumers are identical with respect to preferences and budgets, aggregate consumption,  $C$ , can be expressed as

$$C \equiv Lc = \left( \sum_i y(i)^\gamma \right)^{\frac{1}{\gamma}}. \quad (9.1)$$

Profit maximization leads to prices being markups over the wage rate,  $p(i) = \frac{w}{\gamma b(i)}$ , where  $C$  is the numeraire. The share of labor employed by firm  $i$  and its output are given by

$$\frac{l(i)}{L} = \frac{b(i)^{\frac{\gamma}{1-\gamma}}}{\sum_{i'} b(i')^{\frac{\gamma}{1-\gamma}}} \quad (9.2)$$

$$y(i) = \frac{b(i)^{\frac{1}{1-\gamma}}}{\sum_{i'} b(i')^{\frac{1}{1-\gamma}}} L. \quad (9.3)$$

From this last expression follows that per capita consumption is a function of the  $b_i$  only.

$$c = \left( \sum_i b(i)^{\frac{\gamma}{1-\gamma}} \right)^{\frac{1-\gamma}{\gamma}} \quad (9.4)$$

We have now arrived at the formula that forms the basis for the discussion of the effect of idiosyncratic firm performance on growth. What remains to be done is the specification of the randomness itself. Although it is customary in growth theory to focus on stochastic improvements in quality that result from R&D efforts, I will concentrate on random changes in a firm's labor requirement,  $b$ . The causes of these fluctuations remain unspecified. The choice for shocks in labor requirements simplifies the analysis considerably as random quality improvements would require research, and research, in turn, would only be possible if consumers would save part of their income.

The  $b(i)$  change over time according to a simple Markov process of the form<sup>4</sup>

$$b_t(i) = z_t(i) b_{t-1}(i). \quad (9.5)$$

The stochastic variable  $z > 0$  has realizations that are independent over time and across firms. It has a stable distribution with expected value  $E[z] = 1$  and variance  $\sigma_z^2 < \infty$ . Although for our purposes the  $b(i)$ 's can have any stable distribution at  $t = 0$ , we will impose for all  $i$  that  $b_0(i) = b_0$ , meaning that all firms start with the same productivity.

This Markov process has two properties. First, the expected sample mean (defined as  $\bar{b} \equiv m[b] \equiv 1/n(\sum b_i)$ ) is constant (theorem 7, appendix 9.B.1). If the expected sample mean of  $b$  would not be constant over time, this would have an effect on economic growth that is additional to those caused by Fisher's theorem. Second, the variance of  $b$  increases over time (theorem 8, appendix 9.B.1). We will see that this property enables sustainable growth in the model.

Following the procedure used in the proof of theorem 7, per capita consumption at time  $t$  can be written as a function of the disturbances.

$$c_t = b_0 \left( \sum_i \prod_{\tau=0}^t z_\tau(i)^{\frac{\gamma}{1-\gamma}} \right)^{\frac{1-\gamma}{\gamma}} \quad (9.6)$$

For  $\gamma = 1/2$ , or, an elasticity of substitution of 2, this reduces to  $b_0 \sum_i \prod_{\tau=0}^t z_\tau(i)$  and consequently the expected value of consumption at  $t = 0$  equals  $nb_0$  for any future period. Consumption is therefore not expected to change over time if  $\gamma = 1/2$ .

Unfortunately, if  $\gamma \neq 1/2$ , we cannot derive any further analytical results on the evolution of  $c$  without making additional assumptions. One way to proceed is to assume a continuum of firms with mass  $A$ , such that the consumption function and its reduced form become

$$c_t = \frac{1}{L} \left( \int_0^A y_t(i)^\gamma di \right)^{\frac{1}{\gamma}} \quad (9.7)$$

$$= b_0 \left( \int_0^A \prod_{\tau=0}^t z_\tau(i)^{\frac{\gamma}{1-\gamma}} di \right)^{\frac{1-\gamma}{\gamma}} \quad (9.8)$$

A continuum of firms implies that the randomness in the consumption index disappears, while the mechanism of variation and selection induced by stochastic firm productivity is preserved. Using theorem 9 from the appendix  $c_t$  can be shown to change deterministically over time.

$$c_t = b_0 A^{\frac{1-\gamma}{\gamma}} \mu \left[ z^{\frac{\gamma}{1-\gamma}} \right]^{\frac{1-\gamma}{\gamma} t} \quad (9.9)$$

$\mu$  is an operator for the mean of the population.

Whether consumption is increasing over time depends on  $\mu \left[ z^{\frac{\gamma}{1-\gamma}} \right]$  being larger than one (recall that  $\frac{1-\gamma}{\gamma} > 0$  and  $\mu[z] = 1$ ). If  $\gamma > 1/2$  then  $z^{\frac{\gamma}{1-\gamma}}$  is convex, if  $\gamma < 1/2$  it is concave. In the first case, by Jensen's inequality,

$$\mu \left[ z^{\frac{\gamma}{1-\gamma}} \right] > 1 \quad (9.10)$$

and in the second case

$$\mu \left[ z^{\frac{\gamma}{1-\gamma}} \right] < 1. \quad (9.11)$$

According to theorem 10 (appendix 9.B.1), an increase in the variance of  $z$  makes  $\mu \left[ z^{\frac{\gamma}{1-\gamma}} \right]$  deviate more from one. The size of the permanent shocks to firm performance, measured by the variance of  $z$ , has long term effects on the growth rate of the economy. Only when the elasticity of substitution equals two ( $\gamma = 1/2$ ) this does not apply as in this case  $c$  is a linear function of the  $z(i)$ .

The effect of stochastic firm performance on consumption can be decomposed into two effects. First, the variance in the productivity causes a reallocation of labor towards the more efficient firms. Equation 9.2 shows that a firm with a high  $b$  attracts a larger share of the workforce than a firm with a low  $b$ . Naturally, the quantity produced by the low-cost firm increases even more than employment does, which can be seen from equation 9.3. This shift towards more productive firms leads to a rise in the total (unweighted) quantity produced,  $Y$ .

$$Y \equiv \int_0^A y(i) di = \frac{\int_0^A b(i)^{\frac{1}{1-\gamma}} di}{\int_0^A b(i')^{\frac{\gamma}{1-\gamma}} di'} L = \frac{\mu \left[ b^{\frac{1}{1-\gamma}} \right]}{\mu \left[ b^{\frac{\gamma}{1-\gamma}} \right]} L \quad (9.12)$$

As  $\frac{1}{1-\gamma} > \frac{\gamma}{1-\gamma}$ ,  $b^{\frac{1}{1-\gamma}}$  is more convex than  $b^{\frac{\gamma}{1-\gamma}}$ . Theorem 10 implies that an increase in the variance of  $b$  raises the ratio  $\mu \left[ b^{\frac{1}{1-\gamma}} \right] / \mu \left[ b^{\frac{\gamma}{1-\gamma}} \right]$  and, as a consequence, it raises aggregate production as well.

**Table 9.1.** Simulation setup

		$\sigma_z^2 = 0.01$		
		$\gamma = 0.25$	$\gamma = 0.50$	$\gamma = 0.75$
$n = 100$	sim. 1	sim. 3	sim. 5	sim. 7
figure(s)	9.1	9.1-9.5	9.2-9.5	9.2-9.5
$n = 1000$	sim. 2	sim. 4	sim. 6	sim. 8
figure(s)	9.1	9.1-9.5	9.2-9.5	9.2-9.5

Why is it that the aggregate quantity is positively affected by a rise in the variance of the idiosyncratic shocks for all  $\gamma \in (0, 1)$  while consumption decreases if  $\gamma < 1/2$ ? The answer lies in the loss of the effective variety that is consumed. As time progresses, it becomes increasingly expensive to consume the goods produced with below average productivity while consumers get satiated with the goods that are produced by firms with above average productivity. The mechanism of variation and selection favorable to the aggregate quantity produced, limits the choices of consumers at the same time. Consumers find it increasingly difficult to consume the same amount of every product type.

The aggregate consumption function for a continuum of goods can be written as  $C = (A\mu[y^\gamma])^{\frac{1}{\gamma}}$ . Using this expression the average ‘consumer value’ of a product becomes

$$\frac{C}{Y} = A^{\frac{1-\gamma}{\gamma}} \left( \frac{\mu[y^\gamma]}{\mu[y]^\gamma} \right)^{\frac{1}{\gamma}}. \quad (9.13)$$

As  $y^\gamma$  is concave, Jensen’s inequality implies that  $\mu[y]^\gamma > \mu[y^\gamma]$ . By theorem 10 (appendix 9.B.1), any increase in  $\sigma_b$  reduces the average ‘consumer-value’ of a single unit of a product. This effect is stronger when the substitutability of products is weaker or, equivalently, consumers are satiated more easily and put a high value on product variety. Further discussion of the loss of variety can be found in the next sections.

### 9.3 Small sample results

It is not possible to obtain analytical results for finite samples of firms when  $\gamma \neq 1/2$ . If we want to be assured that the results found for a continuum of firms also apply to finite numbers of firms, we have to rely on Monte-Carlo simulations. Eight different configurations have been run; the results are displayed in figures 9.1 through 9.5 that can be found in appendix 9.A. The size of the workforce is kept constant at 1000 and  $z$  is lognormally distributed with mean one. The other parameters vary per simulation setup and can be read from table 9.1. The number of replications is 300 for each configuration of parameters; the number of periods is 100. Besides the mean (dashed lines) and median (solid lines) results for a configuration, also the 5% and 95% quantiles are plotted, such that 90% of the replications fall between the thin lines.

Figure 9.1 reports the outcomes for the sample variance of  $b$  for  $n \in \{100, 1000\}$  and  $\sigma_z^2 \in \{0.01, 0.05\}$ . This figure does not provide new insights, but is included

as a context for the other simulation results presented below. In all four cases the sample variance of  $b$  increases over time. If  $\sigma_z^2 = 0.05$  then the variance of  $b$  is larger than when  $\sigma_z^2 = 0.01$ , as predicted by theorem 8. The simulations discussed below will all be based on  $\sigma_z^2 = 0.01$ .

Figure 9.2 shows the quantiles for aggregate output,  $Y$ . In all six cases, aggregate output is clearly increasing over time, although the growth rate is higher when products are closer substitutes. A high degree of substitutability between product types lets consumers shift a large part of their expenditure towards low-cost products compared with a situation in which the degree of substitutability is low. Except for the presence of volatility, the behavior of  $Y$  for small samples of firms is identical to the behavior for a continuum of firms described by equation 9.12.

Figure 9.3 displays the evolution of a measure of effective product variety. The adopted measure of product variety is the antilog of Shannon's entropy:

$$N \equiv \prod_i \omega(i)^{-\omega(i)} \quad (9.14)$$

$$\omega(i) \equiv p(i) y(i) / \sum_{i'} p(i') y(i') \quad (9.15)$$

Intuitively,  $N$  can be interpreted as a 'weighted number of products'. The properties of  $N$  are discussed in chapter 5. Product variety is declining for all six configurations but the decline is faster when substitutability is higher. When products are close substitutes, consumers are more willing to concentrate their expenditure on a small range of low-cost products than when products are poor substitutes.

Simulation results for per capita consumption can be found in figure 9.4. The growth rates of the average and median of per capita consumption are negative for  $\varepsilon < 2$ , constant for  $\varepsilon = 2$ , and positive for  $\varepsilon > 2$ . The small sample results are therefore in accordance with the results for a continuum of firms. Figures 9.2 and 9.3 show that idiosyncratic firm performance leads to positive growth in aggregate quantity and negative growth in product variety. As consumers appreciate both the quantity and the variety of consumption goods, these two effects have opposing consequences for consumer welfare. Which of the two effects dominates the other depends on the similarity of product types. When product types are very similar, consumers appreciate quantity more than variety and the net effect of the idiosyncratic shocks on the consumption index is positive. When product types are very dissimilar, variety is better appreciated than quantity and, as a consequence, idiosyncratic shocks will have a negative effect on consumption.

The welfare loss due to reductions in product variety is reflected in the evolution of the average 'consumption value' per single unit of a good. As can be seen from figure 9.5, the average consumption value of a single unit of a good is highest when products are poor substitutes. The consumption value per unit declines for all six cases; the absolute decline being the largest when  $\gamma = 0.25$ . Again, the conclusions drawn for small samples do not deviate from the conclusions drawn for a continuum of firms.

## 9.4 Continuous model

The model of the previous sections can also be formulated in continuous time. As will become clear below, this requires different methods of analysis than have

been used for the discrete time model. Besides being formulated in continuous time, the model of this section also allows for research aimed at the discovery of new products. Because research requires funding, households will have to save part of their income. It is assumed that they will do so by maximizing utility in a standard way.

$$\max_{\{c\}_0^\infty} \left\{ \int_0^\infty \ln [c_t] \exp [-\rho t] dt \right\} \quad (9.16)$$

In this expression,  $c$  is the CES consumption index defined over a continuum of goods stated by equation 9.7. The outcome of the optimization procedure is the familiar Ramsey rule (see section 4.4).

$$\hat{C} = \frac{\pi(A)}{v(A)} - \rho + g_L. \quad (9.17)$$

The optimal growth rate of consumption depends on the profits,  $\pi$ , associated with product type  $A$  relative to the value,  $v$ , of its design or patent.

Like in the discrete time model, the productivity parameter  $b(i)$  is a stochastic variable that is specific to each product and that changes over time. However, unlike the discrete time model, the stochastic process of this variable is an Itô process in stead of a Markov process.<sup>4</sup>

$$b_T = \exp \left[ \ln b_0 + \int_0^T d \ln b_t \right] \quad (9.18)$$

$$d \ln b_t = \left( \alpha - \frac{1}{2} \sigma^2 \right) dt + \sigma dz_t \quad (9.19)$$

$$dz_t = \epsilon_t \sqrt{dt}. \quad (9.20)$$

The firm subscript  $i$  has been omitted in order to simplify notation. The second line is a stochastic differential equation (SDE) with  $\alpha - \frac{1}{2} \sigma_b^2$  being the drift rate,  $dt$  a small increment in time,  $\sigma^2$  the variance of  $d \ln b$ ,  $z_t$  a Wiener process, and  $\epsilon_t$  a disturbance drawn from a stable distribution with mean zero and variance one. The disturbances are independent across firms and time periods.

It is more convenient to work with changes in  $b$  in stead of  $\ln b$ . Applying Itô's lemma to  $db = d(\exp[\ln b])$  shows that  $b$  follows a geometric Brownian motion.

$$b_T = b_0 + \int_0^T db_t \quad (9.21)$$

$$db_t = \alpha b_t dt + \sigma b_t dz_t. \quad (9.22)$$

(Itô's lemma is stated in theorem 11 in appendix 9.B.1.) The expected value of  $b$  is

$$E_0 [b_t] = b_0 \exp [\alpha t]. \quad (9.23)$$

Here,  $E$  is the expected value operator. Note that by defining the drift rate of  $d \ln b$  as  $\alpha - \frac{1}{2} \sigma^2$ , the part of the drift rate that is 'endogenous' to the stochastic process is neutralized.

Just like in the discrete case, the variance of the productivity parameter grows over time, only now the proof is different. The variance of  $b$  is given by  $E[b^2] - (E[b])^2$ . The second part follows immediately from equation 9.23; the first part can be found using Itô's lemma.

$$\text{var}[b_t] = E[b_t^2] - (E[b_t])^2 \quad (9.24)$$

$$= b_0^2 \exp[(2\alpha + \sigma^2)t] - b_0^2 \exp[2\alpha t] \quad (9.25)$$

$$= b_0^2 \exp[2\alpha t] (\exp[\sigma^2 t] - 1) \quad (9.26)$$

From the last expression we can conclude that whenever  $\sigma > 0$ , the variance of  $b$  will grow over time.

The profits made by firms are a markup over their wage bills.

$$\pi(i) = \frac{1-\gamma}{\gamma} w l(i) \quad (9.27)$$

The amount of labor employed by firm  $i$  can be replaced using the counterpart of equation 9.2 for a continuum of firms.

$$\pi(i) = \frac{1-\gamma}{\gamma} w \frac{b(i)^{\frac{\gamma}{1-\gamma}}}{\int_0^A b(i')^{\frac{\gamma}{1-\gamma}} di'} (L - l_n) \quad (9.28)$$

The flow of newly invented product designs simply equals the number of researchers,  $\dot{n} = l_n$ . Entry in the research sector is free, such that the value of a design equals the wage rate,  $v = w$ . It remains to be settled what level of productivity is associated with a new product. According to the Ramsey rule, the ratio of the profits of firm  $A$  to the wage rate should not change over time if the growth rate of consumption is to be constant. From equation 9.28 follows that

$$\frac{\pi(A)}{w} = \frac{(1-\gamma) b(A)^{\frac{\gamma}{1-\gamma}} (L - l_n)}{\gamma \mu [b^{\frac{\gamma}{1-\gamma}}] A} \quad (9.29)$$

should be constant in the steady state. As the growth rate of  $A$  equals the growth rate of  $L - l_n$  in the steady state,  $b(A)^{\frac{\gamma}{1-\gamma}}$  should be a constant multiple of  $\mu [b^{\frac{\gamma}{1-\gamma}}]$ . In other words, the profits of a new firm should be a constant multiple of the average profits. This constant ratio of profit rates is assumed to be one for simplicity. The Ramsey rule can now be stated as

$$\dot{C} = \frac{(1-\gamma)(L - l_n)}{\gamma A} - \rho + g_L. \quad (9.30)$$

The counterpart of equation 9.4 for a continuum of firms is

$$C = \left( \int_0^A b(i)^{\frac{\gamma}{1-\gamma}} di \right)^{\frac{1-\gamma}{\gamma}} (L - l_n) \quad (9.31)$$

The value of  $b(i)^{\frac{\gamma}{1-\gamma}}$  can be derived using Itô's lemma.

$$b_t^{\frac{\gamma}{1-\gamma}} = b_0^{\frac{\gamma}{1-\gamma}} + \int_0^t db_t^{\frac{\gamma}{1-\gamma}} \quad (9.32)$$

$$db_t^{\frac{\gamma}{1-\gamma}} = b_t^{\frac{\gamma}{1-\gamma}} \left( \frac{\gamma\alpha}{1-\gamma} + \frac{\gamma(2\gamma-1)\sigma^2}{2(1-\gamma)^2} \right) dt + b_t^{\frac{\gamma}{1-\gamma}} \frac{\gamma}{1-\gamma} \sigma dz_t \quad (9.33)$$

Application of theorem 9 shows that aggregate consumption behaves deterministically over time.

$$C_t = b_0 \exp \left[ \left( \alpha + \frac{(2\gamma-1)\sigma^2}{2(1-\gamma)} \right) t \right] A_t^{\frac{1-\gamma}{\gamma}} (L_t - l_{n,t}) \quad (9.34)$$

The derivation of this expression can be found in appendix 9.B.2.

The steady state growth rate of consumption,  $g_C$ , depends on the steady state growth rates of the amount of labor employed in the production sector,  $g_L$ , and the number of product types,  $g_A$ , which equals  $g_L$ .

$$g_C = \alpha + \frac{(2\gamma-1)\sigma^2}{2(1-\gamma)} + \frac{1}{\gamma}g_L \quad (9.35)$$

The growth rate of consumption does not only depend on the growth rate of the workforce, but also on the parameters of the stochastic process of  $b$ . That consumption growth depends positively on the drift rate  $\alpha$  is not surprising. The parameter  $\alpha$  is just the exogenous growth rate of expected productivity. More remarkable is the occurrence of  $\sigma$  in the expression above. The effect of the variance of the productivity shocks is positive if  $\varepsilon > 2$  and negative if  $\varepsilon < 2$ . This result is in accordance with the findings for the discrete time model. Only if products are relatively close substitutes, consumers gain more from the rise in aggregate productivity than they lose from the decline in product variety.

The evolution of effective product variety is affected by two forces. First, idiosyncratic shocks have a negative effect on effective product variety – just like in the discrete time model. Second, the number of products grows over time causing effective variety to rise. Effective variety can be measured by a variant of the variety index of equation 9.14 that is suited to a continuum of products.

$$N \equiv \exp \left[ \int_0^A -\omega(i) \ln \omega(i) di \right] \quad (9.36)$$

$$\omega(i) \equiv \frac{p(i)y(i)}{\int_0^A p(i')y(i') di'} \quad (9.37)$$

After some more elaborate manipulation, the influence on  $N$  of both idiosyncratic shocks and the invention of new products becomes clearly visible.

$$N_t = A_t \exp \left[ -\frac{1}{2}\sigma^2 \left( \frac{\gamma}{1-\gamma} \right)^2 t \right] \quad (9.38)$$

The derivation of this result can be found in appendix 9.B.3.

For a constant number of product types, idiosyncratic shocks to the productivity of firms lead to a reduction in effective product variety. This phenomenon was also encountered in discrete time (figure 9.3). As the number of product types grows over time due to R&D efforts, effective variety might also increase. According to equation 9.38 this will only happen if  $g_A = g_L > -\frac{1}{2}\sigma^2 \left( \frac{\gamma}{1-\gamma} \right)^2$ . Without idiosyncratic shocks all firms are symmetric and effective product variety simply equals the number of product types.<sup>5</sup>

## 9.5 Conclusion

Idiosyncratic and permanent shocks to the productivity of firms form a source of economic growth that is distinct from the commonly identified sources of economic growth, which are based on the forward-looking behavior of agents. By inducing growth in the variance of firm performance, permanent idiosyncratic shocks trigger a selection process that leads to a rise in the production of low-cost firms and a decline in the production of high-cost firms. As a consequence, the total unweighted quantity of production increases while effective product variety decreases.



The effect on consumer welfare can be either positive or negative. Consumers appreciate the rise in the quantity of goods they consume but they do not like the loss in variety. The negative effect of the decrease in variety is stronger when products are poor substitutes. In particular, a substitution elasticity with a value below two causes the net effect of idiosyncratic shocks on welfare to be negative.

The mechanism outlined above has been demonstrated to apply to models in discrete time with a discrete or continuous range of firms. The results obtained from discrete time settings generalize to models in continuous time. The continuous time model is extended to allow for the introduction of new goods. A consequence of this feature is that effective product variety may either rise or decline over time, depending on whether the arrival of new goods dominates the concentration effect that is due to idiosyncratic shocks.

The models presented above all predict a negative correlation between product variety and aggregate output. This finding contrasts starkly with the findings of chapters 4, 7, and 8. Some empirical evidence on the relevance of this negative correlation will be presented in chapter 10.

## 9.A Simulation results

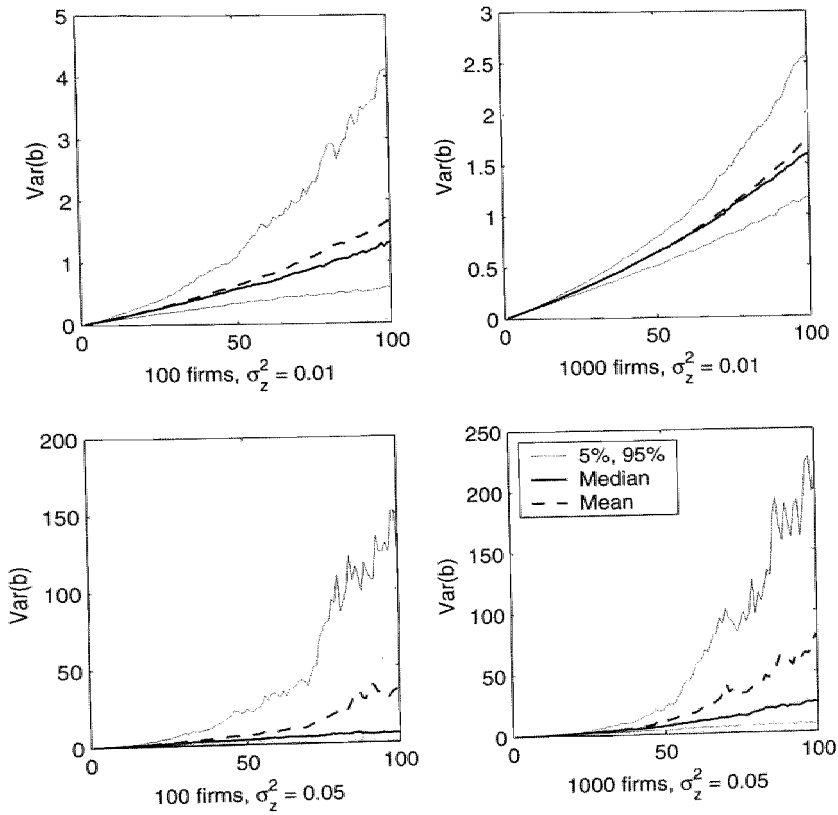


Figure 9.1. Sample variance of  $b(i)$

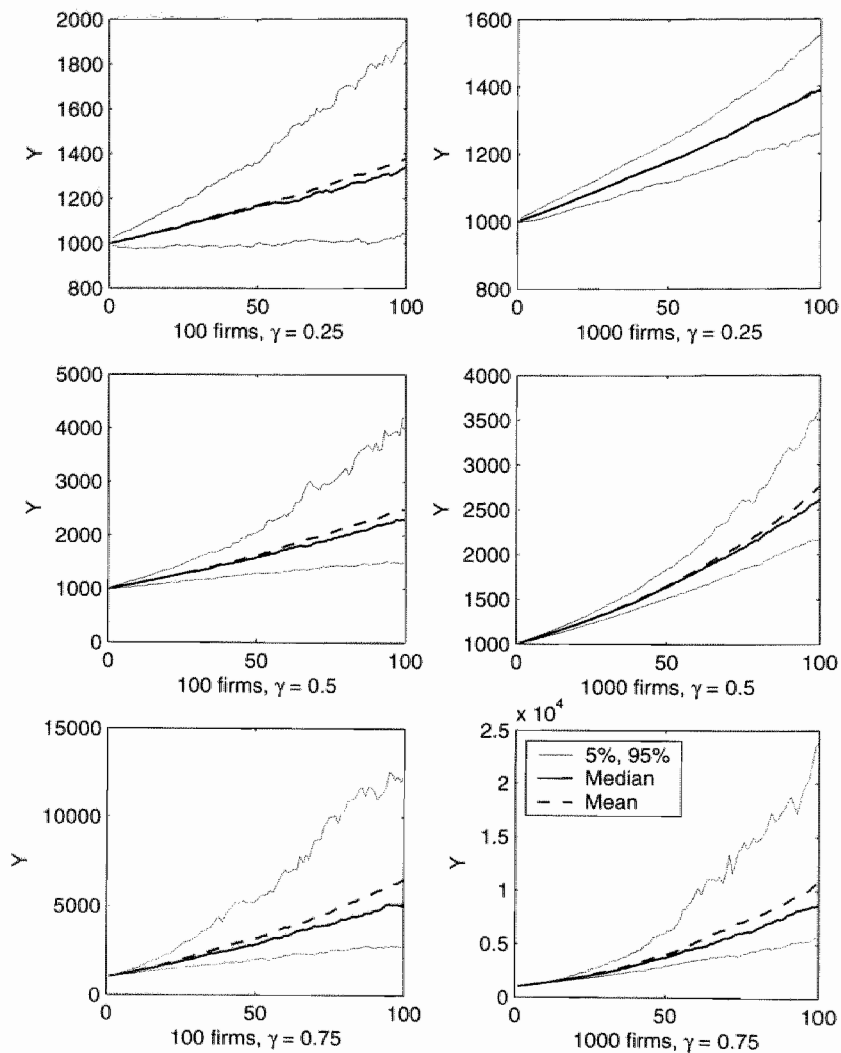
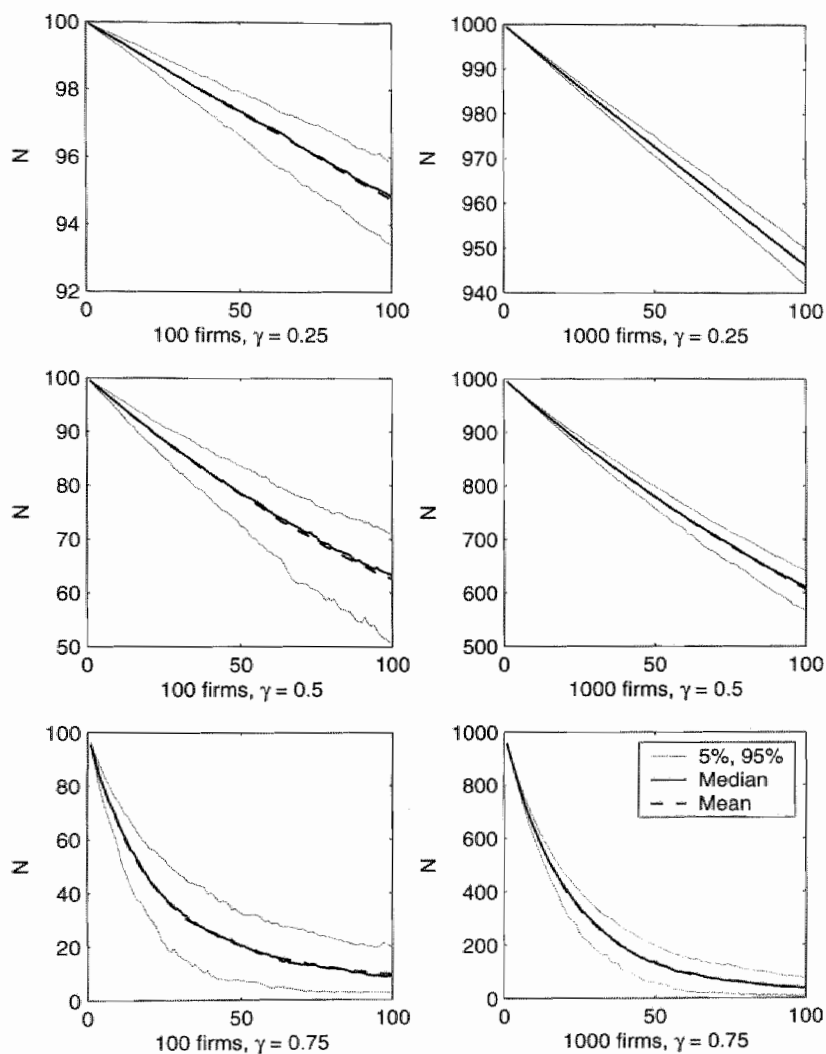
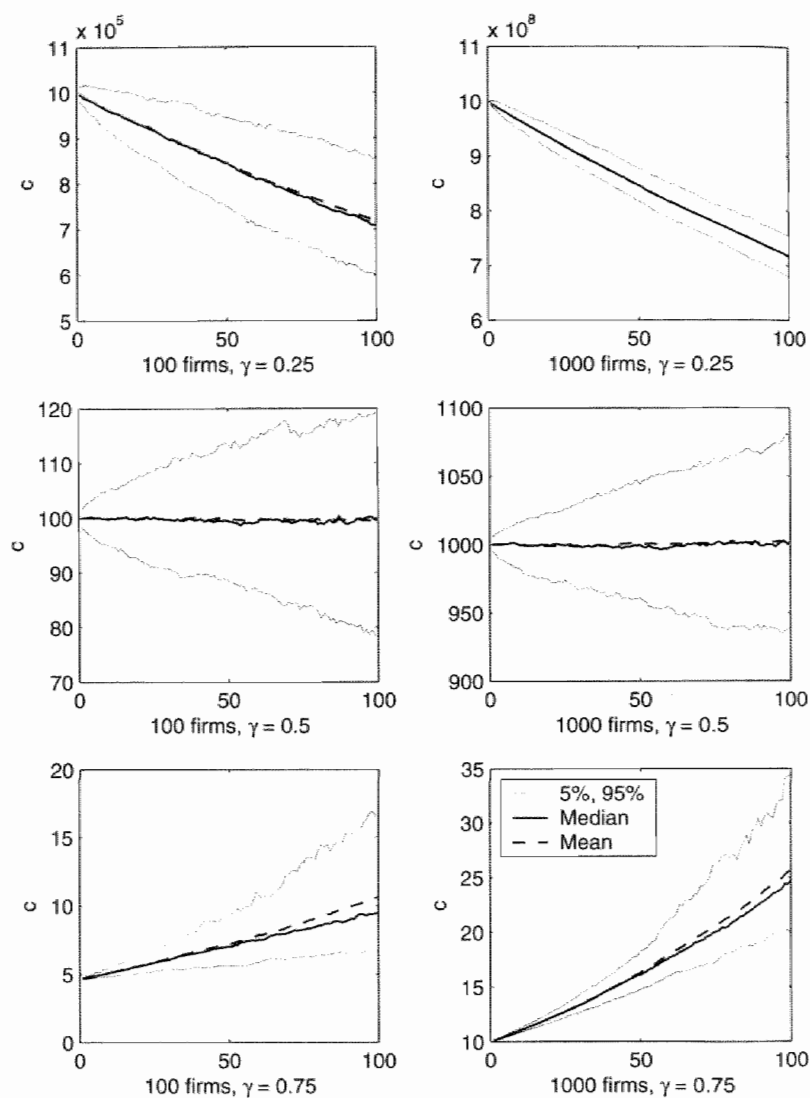


Figure 9.2. Aggregate output ( $Y$ )

**Figure 9.3.** Product variety ( $N$ )



**Figure 9.4.** Consumption per capita ( $c$ )

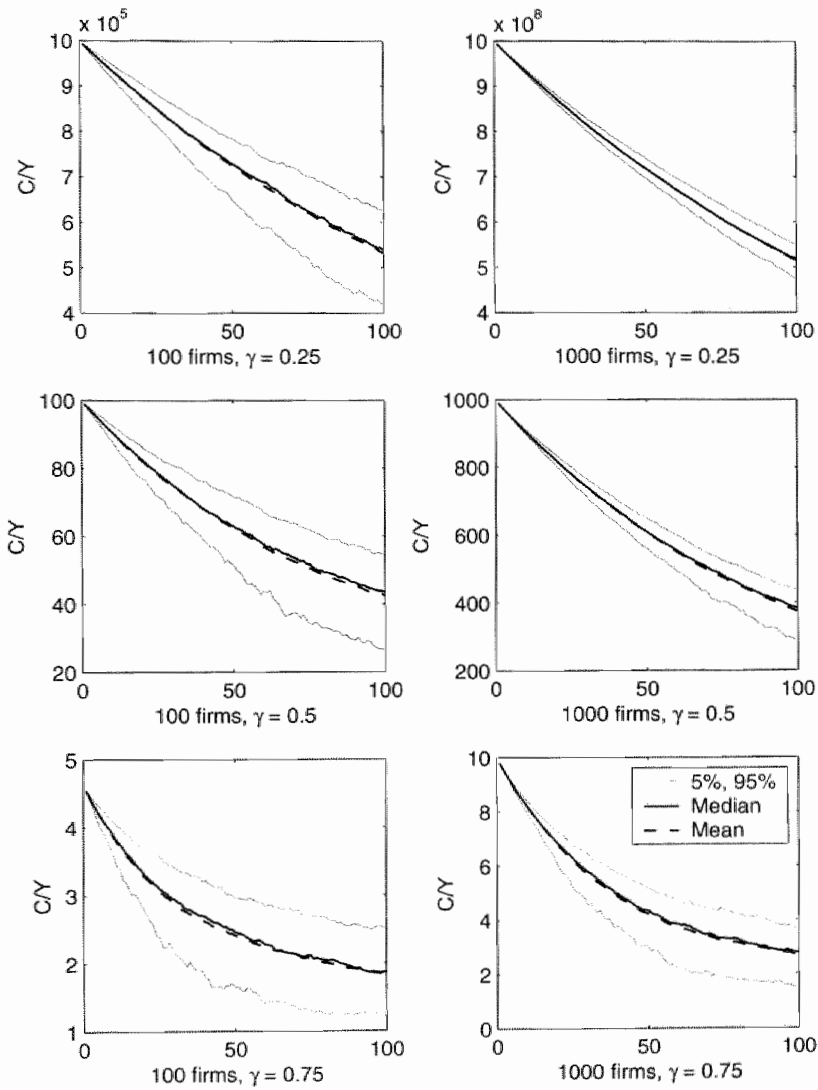


Figure 9.5. Consumption value per unit of production ( $C/Y$ )

## 9.B Mathematical appendix

### 9.B.1 Theorems

**Theorem 7** For a sample of  $n$  stochastic processes of the form  $b_t(i) = z_t(i) b_{t-1}(i)$  with  $z > 0$  i.i.d.,  $E[z] = 1$ , and  $\sigma_z^2 < \infty$ , the expected value of the sample mean at period 0 is constant over time:  $E_0[\bar{b}_t] = \bar{b}_0$ .

**Proof.**

$$E_0[\bar{b}_t] = E_0\left[\frac{1}{n} \sum z_t(i) b_{t-1}(i)\right] = E_0\left[\frac{1}{n} \sum b_0(i) \prod_{\tau=1}^t z_\tau(i)\right] \quad (9.39)$$

$$= \frac{1}{n} \sum b_0(i) E_0\left[\prod_{\tau=1}^t z_\tau(i)\right] = \frac{1}{n} \sum b_0(i) E[z]^t = \bar{b}_0 \quad (9.40)$$

■

**Theorem 8** For stochastic processes of the form  $b_t = z_t b_{t-1}$  with  $z > 0$  i.i.d.,  $E[z] = 1$ , and  $\sigma_z^2 < \infty$ , the variance of  $b$  increases over time according to  $\sigma_{b_t}^2 = ((\sigma_z^2 + 1)^t - 1) b_0^2$ .

**Proof.** The expected sample variance,  $s^2$ , of  $b_t$  at 0 can be decomposed into two parts:

$$E_0[s_{b_t}^2] = E_0[m[b_t^2]] - E_0[\bar{b}_t^2]. \quad (9.41)$$

$m$  is the operator for the sample mean. The first part can be stated as follows

$$E_0[m[b_t^2]] = E_0\left[\frac{1}{n} \sum z_t^2(i) b_{t-1}^2(i)\right] = E_0\left[\frac{1}{n} \sum b_0^2(i) \prod_{\tau=1}^t z_\tau^2(i)\right] \quad (9.42)$$

$$= \frac{1}{n} \sum b_0^2(i) E_0\left[\prod_{\tau=1}^t z_\tau^2(i)\right] = \frac{1}{n} \sum b_0^2(i) (\sigma_z^2 + E[z]^2)^t \quad (9.43)$$

$$= (\sigma_z^2 + E[z]^2)^t m[b_0^2] \quad (9.44)$$

Although for our purposes any sample of  $b$ 's can have any stable distribution at  $t = 0$ , we will impose for any stochastic process  $i$  that  $b_0(i) = b_0$ . The expected sample variance now reduces to  $(\sigma_z^2 + 1)^t b_0^2 - E_0[\bar{b}_t^2]$ . In the limit case where the number of stochastic processes approaches infinity we have  $\lim_{n \rightarrow \infty} E_0[\bar{b}_t^2] = E_0[b_t(i)]^2$  from which it follows that  $\lim_{n \rightarrow \infty} E_0[s_{b_t}^2] = ((\sigma_z^2 + 1)^t - 1) b_0^2$ . ■

**Theorem 9** (Law of Large Numbers for continuum of firms) If  $g$  has a stable distribution then  $\int_0^A g(i) di = A\mu[g]$ , where  $\mu$  is the mean operator.

**Proof.** Suppose that firm  $\nu$  (with  $0 \leq \nu < n$ ) is appointed a fraction  $[i_\nu, i_{\nu+1})$  of the domain  $[0, A]$  such that  $i_{\nu+1} - i_\nu = 1/n$ . Also, suppose that the characteristics of the product produced by firm  $\nu$  depend on  $i_\nu \in [i_\nu, i_{\nu+1})$ . Then, the sum of  $g(i_\nu)$  over all  $n$  firms is

$$\sum_{\nu=0}^n g(i_\nu) (i_{\nu+1} - i_\nu) = \sum_{\nu=0}^n g(i_\nu) / n \quad (9.45)$$

The limit of this sum for  $n \rightarrow \infty$  is  $\int_0^A g(i) di$ . If the  $g(i_\nu)$  are realizations of the same stable distribution then by the Law of Large Numbers

$$\text{plim}_{n \rightarrow \infty} \sum_{\nu=0}^n g(i_\nu) / n = A\mu[g]. \quad (9.46)$$

Combining both results yields  $\int_0^A g_i di = A\mu[g]$ . See also Judd (1985a). ■

**Theorem 10** (*Jensen's inequality with changing variance*) If  $h(b)$  is a strictly convex function and  $b$  is a (possibly random) variable, then

$$\mu[h(b) | \sigma_b + d\sigma_b] > \mu[h(b) | \sigma_b] \text{ for } d\sigma_b > 0, \quad (9.47)$$

where  $\mu$  is the mean operator,  $\sigma_b$  is the standard deviation of  $b$  and  $d\sigma_b$  is a (small) increment in the standard deviation of  $b$ . If  $h(b)$  is a strictly concave function, then

$$\mu[h(b) | \sigma_b + d\sigma_b] < \mu[h(b) | \sigma_b] \text{ for } d\sigma_b > 0. \quad (9.48)$$

**Proof.** (First part is identical to Davidson, 1994, lemma 9.13.) Let  $h(b)$  be a strictly convex function. The definition of a convex function ensures that

$$h((1-\lambda)\bar{b} + \lambda b) < (1-\lambda)h(\bar{b}) + \lambda h(b) \quad (9.49)$$

$$\frac{h((1-\lambda)\bar{b} + \lambda b) - h(\bar{b})}{\lambda} < h(b) - h(\bar{b}) \quad (9.50)$$

for  $\lambda \in (0, 1]$ . If  $b \geq \bar{b}$  then

$$\frac{h(\bar{b} + \lambda(b - \bar{b})) - h(\bar{b})}{\lambda(b - \bar{b})} \leq \frac{h(b) - h(\bar{b})}{b - \bar{b}} \quad (9.51)$$

By letting  $\lambda \downarrow 0$ , the right hand side of the expression above becomes the derivative of  $h$ .

$$h'(\bar{b}) \leq \frac{h(b) - h(\bar{b})}{b - \bar{b}} \quad (9.52)$$

Multiplication by  $b - \bar{b}$ , squaring, and taking means yields

$$[h'(\bar{b})]^2 (b - \bar{b})^2 < [h(b) - h(\bar{b})]^2 \quad (9.53)$$

$$\mu \left[ [h'(\bar{b})]^2 (b - \bar{b})^2 \right] < \mu \left[ [h(b) - h(\bar{b})]^2 \right] \quad (9.54)$$

$$[h'(\bar{b})]^2 \sigma_b^2 < [\mu[h(b)] - h(\bar{b})]^2 \quad (9.55)$$

where  $\sigma_b^2$  is the variance of  $b$ . By Jensen's inequality and  $h$  being convex,  $\mu[h(b)] > h(\bar{b})$  and thus

$$\sigma_b < \frac{\mu[h(b)] - h(\bar{b})}{|h'(\bar{b})|}. \quad (9.56)$$

As  $\bar{b}$  is fixed,  $d\sigma_b < d\mu[h(b)] / |h'(\bar{b})|$  and

$$\frac{\mu[h(b)] - h(\bar{b})}{|h'(\bar{b})|} < \frac{\mu[h(b)] + d\mu[h(b)] - h(\bar{b})}{|h'(\bar{b})|}. \quad (9.57)$$



Multiply by  $|h'(\bar{b})|$  and add  $h(\bar{b})$  on both sides to prove the theorem.

$$\mu[h(b)|\sigma_b + d\sigma_b] > \mu[h(b)|\sigma_b] \text{ for } d\sigma_b > 0. \quad (9.58)$$

If  $h$  would be a concave function, we would have

$$\mu[h(b)|\sigma_b + d\sigma_b] < \mu[h(b)|\sigma_b] \text{ for } d\sigma_b > 0. \quad (9.59)$$

■

**Theorem 11** (*Itô's lemma*) Consider the Itô process  $dx = \alpha dt + \sigma dz_t$  with  $dz = \epsilon_t \sqrt{dt}$  being a Wiener process and let  $h(x)$  be a continuous function of  $x$ . Then  $dh(x) = [h'\alpha + \frac{1}{2}h''\sigma^2]dt + h'\sigma dz$ .

**Proof.** A formal proof can be found in, for example, Malliaris and Brock (1982, pp. 80-92). Below a sketch of the proof is given. First approximate  $h(x)$  with a second order Taylor expansion and substitute for  $dx$ .

$$dh(x) = h'dx + \frac{1}{2}h''(dx)^2 \quad (9.60)$$

$$= h'(\alpha dt + \sigma dz) + \frac{1}{2}h''(\alpha dt + \sigma dz)^2 \quad (9.61)$$

Second, apply the rules  $(dt)^2 = 0$  and  $(dz)^2 = dt$  to get

$$dh(x) = h'\alpha dt + h'\sigma dz + \frac{1}{2}h''\sigma^2 dt. \quad (9.62)$$

■

## 9.B.2 Derivation of $C$ in continuous time

The main problem with the derivation of equation 9.34 is the integral in equation 9.31, which is labeled  $B$  below. Decompose  $B$  into an integral for all product types present at  $t = 0$  and an integral for the product types invented after  $t = 0$ .

$$B_T \equiv \int_0^{A_T} b_T(i)^{\frac{\gamma}{1-\gamma}} di = B_0^{A_0} + B_{A_0}^{A_T} \quad (9.63)$$

$$B_0^{A_0} \equiv \int_{i=0}^{A_0} \left( b_0^{\frac{\gamma}{1-\gamma}} + \int_{t=0}^T db(i)_t^{\frac{\gamma}{1-\gamma}} \right) di \quad (9.64)$$

$$B_{A_0}^{A_T} \equiv \int_{\tau=0}^T \int_{i=A_\tau}^{A_{\tau+d\tau}} \left( b_\tau^{\frac{\gamma}{1-\gamma}} + \int_{t=\tau}^T db_t(i)^{\frac{\gamma}{1-\gamma}} \right) di d\tau \quad (9.65)$$

The initial profitability of a good invented at time  $\tau$  equals the average profitability of the existing products. Using theorem 9 this implies

$$b_\tau^{\frac{\gamma}{1-\gamma}} = \frac{1}{A_\tau} \int_0^{A_\tau} b_\tau(i)^{\frac{\gamma}{1-\gamma}} di = b_0^{\frac{\gamma}{1-\gamma}} \exp \left[ \left( \frac{\gamma\alpha}{1-\gamma} + \frac{\gamma(2\gamma-1)\sigma^2}{2(1-\gamma)^2} \right) \tau \right]. \quad (9.66)$$

Substitute for  $b_\tau^{\frac{\gamma}{1-\gamma}}$  in the integral for products invented after  $t = 0$  and apply theorem 9 once more.

$$B_{A_0}^{A_T} = \int_{\tau=0}^T b_\tau^{\frac{\gamma}{1-\gamma}} \exp \left[ \left( \frac{\gamma\alpha}{1-\gamma} + \frac{\gamma(2\gamma-1)\sigma^2}{2(1-\gamma)^2} \right) (T-\tau) \right] (A_{\tau+d\tau} - A_\tau) d\tau \quad (9.67)$$

$$= \int_{\tau=0}^T b_0^{\frac{\gamma}{1-\gamma}} \exp \left[ \left( \frac{\gamma\alpha}{1-\gamma} + \frac{\gamma(2\gamma-1)\sigma^2}{2(1-\gamma)^2} \right) T \right] (A_{\tau+d\tau} - A_\tau) d\tau \quad (9.68)$$

$$= b_0^{\frac{\gamma}{1-\gamma}} \exp \left[ \left( \frac{\gamma\alpha}{1-\gamma} + \frac{\gamma(2\gamma-1)\sigma^2}{2(1-\gamma)^2} \right) T \right] (A_T - A_0) \quad (9.69)$$

For products in ranging from 0 to  $A_0$ , the integral is given by

$$B_0^{A_0} = b_0^{\frac{\gamma}{1-\gamma}} \exp \left[ \left( \frac{\gamma\alpha}{1-\gamma} + \frac{\gamma(2\gamma-1)\sigma^2}{2(1-\gamma)^2} \right) T \right] A_0. \quad (9.70)$$

Adding up  $B_0^{A_0}$  and  $B_{A_0}^{A_T}$  yields

$$B_T = b_0^{\frac{\gamma}{1-\gamma}} \exp \left[ \left( \frac{\gamma\alpha}{1-\gamma} + \frac{\gamma(2\gamma-1)\sigma^2}{2(1-\gamma)^2} \right) T \right] A_T. \quad (9.71)$$

Insert the solution for  $B$  into equation 9.31 to get equation 9.34.

### 9.B.3 Derivation of $N$ in continuous time

Define  $B \equiv \int_0^A b(i')^k di'$  and  $k \equiv \frac{\gamma}{1-\gamma}$ , then the variety index in 9.36 can be written as

$$N = \exp \left[ \int_0^A -\frac{b(i)^k}{B} \ln \left[ \frac{b(i)^k}{B} \right] di \right] \quad (9.72)$$

$$= \exp \left[ \frac{-1}{B} \left( \int_0^A b(i)^k \ln [b(i)^k] di - \ln [B] \int_0^A b(i)^k di \right) \right] \quad (9.73)$$

$$= B \exp \left[ \frac{-1}{B} \int_0^A b(i)^k \ln [b(i)^k] di \right] \quad (9.74)$$

From Itô's lemma and theorem 9 follows that  $B = b_0^k \exp [K_1 t] A_t$  where  $K_1 \equiv (\alpha + \frac{1}{2}\sigma^2(k-1))k$ . What remains to be found is a solution for  $\int_0^A b(i)^k \ln [b(i)^k] di$ .

Theorem 9 implies that

$$\int_0^A b(i)^k \ln [b(i)^k] di = A E[f], \quad f(i) \equiv b(i)^k \ln [b(i)^k]. \quad (9.75)$$

Apply Itô's lemma to get the increment of  $f$  (the index is dropped for notational convenience).

$$df = \left( kb^{k-1}(k \ln [b] + 1) \alpha b + \frac{1}{2} kb^{k-2} [(k-1)(k \ln [b] + 1) + k] \sigma^2 b^2 \right) dt \\ + kb^{k-1}(k \ln [b] + 1) \sigma b_t dz_t \quad (9.76)$$

Take the expected value of  $df$  at period zero, divide by  $dt$ , replace  $E_0 [b_t^k \ln b_t^k]$  by  $E_0 [f_t]$ , substitute for  $E [b_t^k]$ , and rearrange.

$$\frac{E_0 [df]}{dt} = K_1 E_0 [f_t] + K_2 \exp [K_1 t], \quad K_2 \equiv (\alpha + \frac{1}{2}\sigma^2(2k-1)) kb_0^k \quad (9.77)$$

This differential equation in  $E_0 [f]$  has initial condition  $f_0 = b_0^k \ln [b_0^k]$ . After a straightforward application of the procedure of trial and error the solution turns out to be

$$E_0 [f_t] = \exp [K_1 t] (K_2 t + f_0). \quad (9.78)$$

Combine this solution for  $E [f]$  with equations 9.74 and 9.75 and substitute for  $B$  and  $K_2$  in the exponent.

$$N_t = B_t \exp \left[ - \left( (\alpha + \frac{1}{2}\sigma^2(2k-1)) kt + \frac{f_0}{b_0^k} \right) \right]. \quad (9.79)$$

Equation 9.38 is found after substitution of  $B$ ,  $k$ , and  $f_0$ .

## Notes

1. The mechanism described in this chapter differs from the one presented by Fisher as the models of this chapter are not genetical. Nevertheless, the abstract principle described by Fisher's theorem does apply to these models.
2. For detailed evidence on the volatility of a firm's profit and investment see Scherer (1998), Scherer, Harhoff, and Kukies (2000), and Schankerman (2002).
3. The 'shotnoise' processes used by e.g. Dinopoulos and Thompson (1998), and Howitt (1999) lead to convergence of the distribution of *relative* productivity parameters to a stationary distribution, but it does not lead to a stationary distribution of the *actual* productivity parameters. In contrast, Thompson (2001) arrives at an asymptotically stationary distribution of firm size. This implies that in his model the effects of Fisher's theorem vanish asymptotically.
4. No drift has been included in order to keep the model as simple as possible. A stochastic process has a drift when positive shocks are of a different size than negative shocks. Straathof (2002) shows that including a drift only has trivial effects on the growth rate of an economy.
5. Asymmetry in firms is also present in the model of section 7.3 where new products are imperfect improvements over older products. However, the entropy of product sales in that model remains constant over time (equation 7.66) and therefore the asymmetry of firms has no effects on the growth of welfare.

## Empirical analysis

In this chapter the theoretical considerations of the foregoing chapters are confronted with data on product variety. Both a cross-sectional and a time-series analysis are conducted. The cross-sectional analysis suggests that the scale of production is an important determinant of the number of produced product types in small countries, while produced variety in large countries seems to be constrained by the worldwide stock of inventions. The time-series analysis shows that within countries, product variety can largely be explained by the stock of human capital. Furthermore, the ratio of effective variety to the number of types is found to be declining in most countries, which points to the presence of idiosyncratic industry performance.

### 10.1 Introduction

The foregoing chapters discussed various perspectives on the relation between product variety and consumer welfare. If one is to summarize this relation, then three factors constitute to the theoretical discussion: scale, knowledge spillovers, and idiosyncratic firm performance. That the scale of an economy affects the number of product types is disputed neither in the foregoing chapters, nor in the theoretical literature as a whole. Whether knowledge spillovers may also serve as an autonomous determinant of product variety or whether they merely amplify the effects of scale, has been the subject of debate. Judging from the more recent neo-endogenous growth models, however, a consensus seems to have been reached in support of the latter position that variety ultimately depends on scale. (Many of the models in this dissertation even abstract from knowledge spillovers altogether.)

Idiosyncratic firm performance, the third factor, has an impact on effective variety – a notion discussed in chapter 5 – rather than on the number of products. Chapter 9 shows how idiosyncratic productivity shocks can simultaneously reduce effective variety and raise consumer welfare. When it comes to effective variety, knowledge spillovers also play a role. In the second model of chapter 7, spillovers cause effective variety to be lower than the number of products because new products are imperfect improvements of old products.

The aim of the current chapter is to shed some light on the realism of all these theoretical considerations. In particular, evidence is presented that confirms the existence of a relation between scale and the number of products. In addition, evidence has been found on the relation between idiosyncratic firm performance

and effective product variety. The overall picture that arises is that knowledge spillovers are only important as a geographical phenomenon: the production of a product type clearly is not limited to the country in which it is invented. No indication has been found that intertemporal knowledge spillovers matter for the number of product types.

The empirical literature on product variety is rather small. There are two likely causes for a lack of enthusiasm to work on this topic: a lack of data and a lack of theory. The first reason for discouragement is due to the limited availability of standardized, detailed production data. A commonly used remedy against this problem is to use export data instead. Another data related issue is the possibility of bias stemming from the use of product groups rather than actual product types.<sup>1</sup>

Besides the practical problem of data availability, there is also the theoretical problem of how product variety should affect welfare if not all types are of equal importance. To date, nearly all empirical work on product variety is based on the theoretical work of Feenstra (Feenstra, Markusen, and Zeile 1992, Feenstra 1994, Feenstra and Markusen 1994, Feenstra and Kee 2004). A discussion of Feenstra's variety index can be found in section 6.D. Feenstra's index has been applied to trade data by amongst others Feenstra, Madani, Yang, and Liang (1999), Feenstra, Yang, and Hamilton (1999), Funke and Ruhwedel (2001a, 2001b, 2002), Hummels and Klenow (2002), and Broda and Weinstein (2004a, 2004b). There are two reasons for questioning the reliability of these studies. First, the Feenstra index is not exact and is only defined for CES function that are homogenous of degree one. Second, breaks in the number of products due to classification changes are not taken into account.<sup>2</sup>

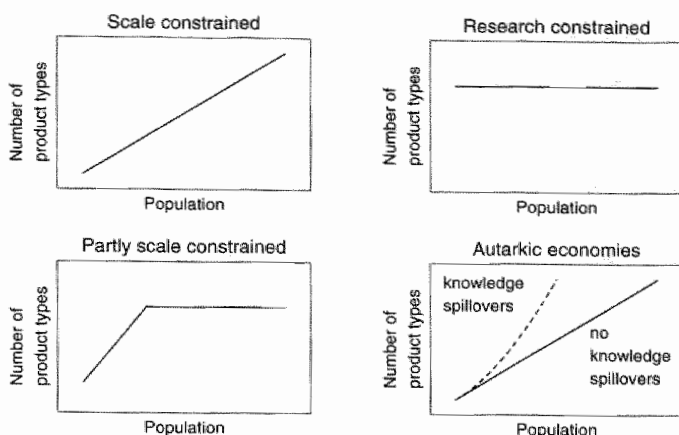
Taking a different approach than Feenstra, Bils and Klenow (2001) find evidence of an acceleration in the growth rate of product variety. They use data on household consumption from the US Bureau of Labor Statistics (BLS) to show that the expenditure share of product groups characterized by a high rate of variety change has grown substantially between 1980 and 1999 (see also Klenow 2003).

Inspired by the works of Hicks (1940), Hausman (1981, 1997) has developed a methodology for estimating the welfare gains that arise from the introduction of a new good. Later contributions in this direction include the work by Petrin (2001), Goolsbee and Petrin (2001), and Nevo (2003). Authors that have studied the effects of new goods using alternative approaches are Trajtenberg (1989), Brynjolfsson and Hitt (2003), and Berry, Levinsohn, and Pakes (2004). All of the latter studies are partial equilibrium in nature and are not concerned with product variety as such, but they do provide some indirect evidence that product variety is appreciated by consumers.

The remainder of this chapter is organized as follows. Section 10.2 contains a cross-section analysis of the relation between the number of different goods produced in a country and the size of the country. Section 10.3 treats the bias that might arise from using data on product groups rather than data on actual product types. A longitudinal analysis of the influence of scale on variety can be found in section 10.4, while section 10.5 treats the evolution of effective product variety. The results are summarized and overall conclusions are drawn in section 10.6.

## 10.2 Product variety and country size

According to the theory of chapter 8, the number of types of goods produced by a country essentially depends upon two things: the country's interaction with other



**Figure 10.1.** Theoretical overview of the relation between the number of produced types and population size for a cross section of countries

countries and the size of its population. The interaction between countries may influence the variety of domestically produced goods in two ways. First, economies of scale at the plant or industry level may induce international specialization. In this situation a country's produced variety will depend on the size of its population as the scale of production is determined by the integrated multi-country economy. This is illustrated in the upper-left panel of figure 10.1.

Second, when trade is costly or impossible countries still can interact by imitating each other's inventions. In case that economies of scale are negligible, produced variety will be constrained only by the worldwide stock of inventions. As a result of this, produced variety will be independent of population size in the cross-sectional dimension (upper-right panel).<sup>3</sup> Alternatively, if economies of scale are significant but moderate, then smaller countries will only produce a subset of the types that have been invented worldwide, while large countries will produce all types (lower-left panel).

The situation in which countries do not interact with each other – they neither trade nor imitate – is illustrated by the lower-right panel of figure 10.1. The size of a country now determines the stock of inventions that the country has access to. If intertemporal knowledge spillovers are absent, then the relation between variety and size is linear – just as in the upper-left panel. When spillovers do occur, then the relation becomes convex: product variety per capita is no longer constant but depends on the size of the population (see equation 4.5).<sup>4</sup>

In order to be able to assess the accuracy of the theoretical predictions illustrated by figure 10.1, we need data on the number of types of goods that a country produces. To my knowledge there exist only two sources that provide data that are suitable for cross-sectional analysis: the Economic Census by the US Census Bureau and the EUROPROMS database by Statistics Europe. A description of the data stemming from these sources can be found in appendix A. The data refer to a mixture of final goods and intermediates. Although the theories outlined above are primarily concerned with consumer goods, no attempt is made in this chapter to separate the effects of variety in consumer goods from those of intermediate variety. Thus, it is implicitly assumed that intermediate variety affects the standard of living in roughly the same way as consumer good variety – the

**Table 10.1.** Regression of the log of the number of product groups on the log of population size

	$\ln(P)$	constant	$R^2$			obs. groups
			within	between	overall	
US states <sup>†</sup>	0.226	2.666			0.86	51
	0.018**	0.278**				1
European countries <sup>‡</sup>	0.480	-0.548	0.53	0.07	0.52	90
	0.049**	0.795				6

<sup>†</sup> OLS regression with heteroscedasticity robust covariance

<sup>‡</sup> Panel regression with fixed effects for the years; Luxembourg excluded

\*\* Coefficient significant at the 1% confidence level

Standard deviations below coefficients

only difference being that the gains from intermediate variety reach the consumer through cheaper or better consumer goods.

Figure 10.2 displays a scatter diagram of produced variety and population size for 51 states of the United States. The relation between the number of product groups and population size is clearly concave. The dashed curve is obtained by a linear regression of the log of the number of product groups on the log of the population size. The corresponding estimation results are displayed in the first row of table 10.1. The coefficient on the log of population,  $P$ , is substantially below one, confirming the concavity of the relation.

Comparing the panels in figure 10.1 with figure 10.2, it seems to be that the lower-left panel gives the most accurate description of the data. Assuming that the states of the United States together form a highly integrated economy, the inter-state differences in produced variety should be a consequence of economies of scale at the plant or industry level.

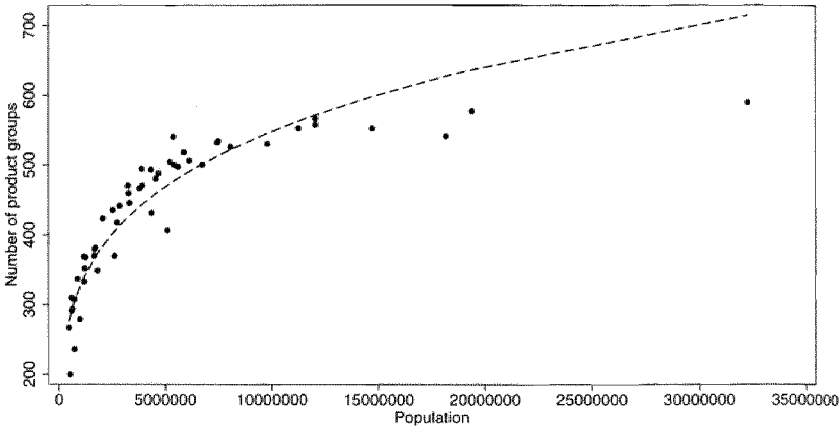
The presence of economies of scale is confirmed by figure 10.3, which shows that the relation between the number of establishments (producing units) and population size is linear. The patterns visible in figures 10.2 and 10.3 together imply that although larger states tend to produce a greater variety of products, they also tend to have a larger number of establishments per industry. The theoretical explanation that suits this observation best is that product variety in the smaller states is constrained by the nation-wide optimum scale of production, while variety in the larger states is constrained by the number of types that has been invented.

The data on European countries do not exhibit the clear patterns that characterize the data on US states. Figure 10.4 contains a scatter diagram of the number of product groups and population size for the period 1995 to 2000. There is still a clear positive relation between variety and size in the cross-sectional dimension, but the concavity is less pronounced. This is confirmed by the estimation results reported in the second row of table 10.1. The coefficient on the log of the population size is closer to one for the European data than it is for the US data. This can be explained by the theory in two ways.<sup>5</sup> First, European countries might be constrained by the optimal scale of production rather than by the number of inventions. Alternatively, Europe's economies could be more reliant on domestic research than the US states are. Remarkable is that Denmark and Sweden appear to have an inclination to imitate rather than trade considering their relatively high levels of produced variety.

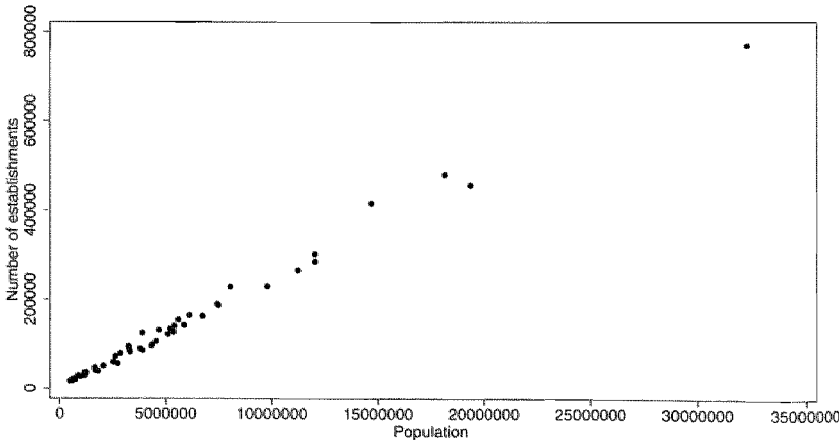
Turning to figure 10.5, the number of establishments appears to be linearly associated with population size for European countries as well, although Italy obviously does not fit into this pattern. This might be due to the fact that countries use different definitions of what constitutes an 'establishment'. In general, it is not surprising that the patterns in the data are less clear for European countries than they are for US states: the economic policies and customs (and statistical agencies!) of these countries are still far from homogeneous.

Cross-section analysis shows that small countries or states tend to produce a smaller number of products than large countries. This result is particularly strong for the US states. For the European countries the pattern is less clear, in particular because of the large variety of production in Denmark and Sweden. The concave association between variety and size within the United States suggests that variety in the smaller states is constrained by the scale of production, whereas in the larger states, variety is constrained by the stock of inventions.<sup>6</sup> No clear conclusions can be drawn for the European countries, except that the absence of clear patterns in the European data seem to suggest that the European economy is much less integrated than the US economy.

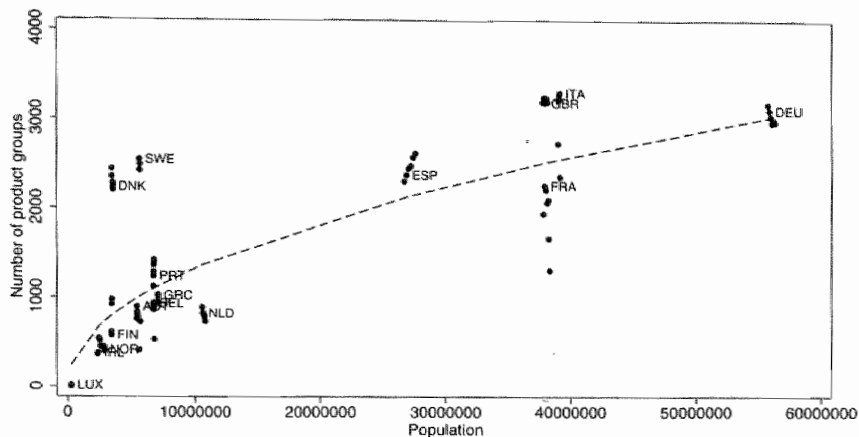




**Figure 10.2.** Number of product groups and population size for US states (1997)  
Source: US Census Bureau

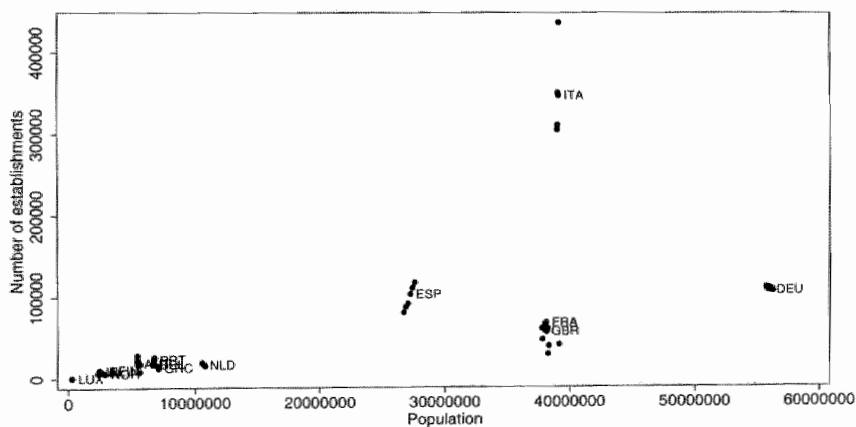


**Figure 10.3.** Number of establishments and population size (US states, 1997)  
Source: US Census Bureau



**Figure 10.4.** Number of product groups and population size for European countries (1995-2000; fitted values for 1997)

Sources: Statistics Europe (EUROPROMS); World Bank (WDI)



**Figure 10.5.** Number of establishments\* and population size (European countries, 1995-2000)

Source: Statistics Europe (EUROPROMS)

\*Definition differs across countries

### 10.3 The curse of classification?

The concave association between product variety and population size of figure 10.2 might be a spurious phenomenon. The process of collecting and reporting data itself may cause concavity in the observed relationship. The reason is that the maximum number of product groups a classification has for a given level of detail, largely depends on arbitrary decisions made by the designers of the classification system.

I will illustrate how these arbitrary choices can lead to concavity with the following theoretical experiment. A certain number of  $n$  product types is to be assigned to at most  $G$  product groups. It is possible that  $n > G$ , such that a group might contain several types. The number of groups that have at least one type is  $g$ . (I will refer to these groups as 'filled' groups.) Suppose that each product type is assigned randomly to a group, regardless of whether any types were assigned to the group previously. Then it is possible to calculate which value of  $n$  has the highest probability of outcome  $g$ . The probability of  $g$  given  $n$  and  $G$  is given by

$$p(g|n, G) = \frac{S_n^{(g)}}{\sum_{h=1}^G S_n^{(h)}}. \quad (10.1)$$

Here,  $S_n^{(g)}$  is the number of ways  $n$  types can be assigned to  $g$  groups, regardless of the order by which types are assigned or the ordering of the groups.<sup>7</sup>  $S_n^{(g)}$  is a Stirling number of the second kind (see Abramowitz and Stegun 1972 p. 824 and Wolfram Research 2004). Its primary definition is stated below.

$$S_n^{(g)} \equiv \frac{1}{g!} \sum_{h=0}^G (-1)^{g-h} \binom{g}{h} h^n \quad (10.2)$$

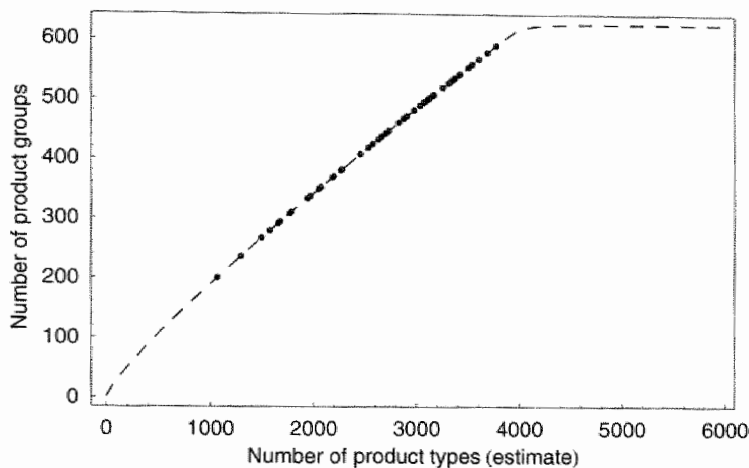
In principle, it is possible to find the  $n$  with the highest probability of  $g$  by increasing  $n$  until  $p(g|n+1) < p(g|n)$  and the maximum has been reached.<sup>8</sup> However, this method requires a large number of computations because of the large binomial coefficients. Lowering the precision of  $n$  is one strategy for reducing computing time, but if  $n$  is large this approach is still very demanding as  $p(g|n, G)$  still has to be computed several times for each  $g$ .

In order to radically reduce computing time, a radical method of approximation is needed. Suppose that the number of actual product types is known, then the expected number of filled groups can be computed relatively quickly. The expected value of  $g$  can be computed in the usual way.

$$E(g|n, G) = \frac{\sum_{g=1}^G S_n^{(g)} g}{\sum_{h=1}^G S_n^{(h)}} \quad (10.3)$$

After having computed the expected value of  $g$  for several values of  $n$ , we can approximate the relation between  $E(g)$  and  $n$  by interpolation of the known points. In reality, of course, we do not know the actual number of types – we only know the actual number of filled groups. A radical approximation that saves a large number amount of computations is based on the assumption that the value of  $n$  that is obtained by setting  $g = E(g|n, G)$  is approximately equal to the value of  $n$  that maximizes  $p(g|n, G)$ .

$$n|_{g=E(g|n, G)} \simeq \arg \max_n \{p(g|n, G)\} \quad (10.4)$$



**Figure 10.6.** Estimated number of product types compared to actual number of product groups (US states, 1997)

Sources: US Census Bureau; author's calculations

Simulations with small values of  $g$ ,  $G$ , and  $n$  suggest that this bold assumption is reasonably accurate.

Figure 10.6 shows the relation between the number of types and the expected number of filled groups for  $G = 626$  – the number of industries in the North American Industry Classification System (NAICS), which is used for the US Economic Census. The expected value of  $g$  has been computed for 100 values of  $n$ , which have been interpolated linearly.<sup>9</sup> The dots represent the number of filled product groups for each of the 51 states reported by the US Economic Census.<sup>10</sup>

The simulated relation between  $n$  and  $g$  is surprisingly linear in the range of the empirical observations. It seems to be that the concavity visible in figure 10.2 cannot be attributed to the grouping of product types – as long as it is realistic to assume that types are assigned randomly and independently to groups. This assumption of random assignment, however, is not likely to hold in reality. A country producing a certain type of goods probably will also produce a range of similar types as well. For example, a country producing luxury cars, like Germany or Italy, is likely to produce other kinds of cars as well. This implies that once a certain product group is 'filled' for a country, other groups will also be filled for that country. If product types are geographically clustered, then so will be product groups.

The presence of clustering brings with it that the assignment of a type is no longer independent of the assignments of previous types, which invalidates the estimates presented in figure 10.6. In order to assess the seriousness of this problem, let us conduct the following extreme experiment. Suppose that both product types and product groups come in clusters of 62.6 elements. (62.6 because then the 626 industries of NAICS would be distributed over ten group-clusters.) If the first product type of type-cluster  $i$  is assigned to a group belonging to group-cluster  $I$ , then automatically the 61.6 other groups of cluster  $I$ , will be filled with the other 61.6 types of cluster  $i$ . It is explicitly assumed that all the types in a type-cluster can only be assigned to groups belonging to the same group-cluster.

**Table 10.2.** Regression of the log of the estimated number of product types on the log of population size

	$\ln(P)$	constant	$R^2$	obs.
Without clustering	0.263 0.021**	3.910 0.326**	0.86	51
With clustering	0.330 0.022**	1.960 0.334**	0.89	51

OLS regression with heteroscedasticity robust covariance

\*\* Coefficient significant at the 1% confidence level

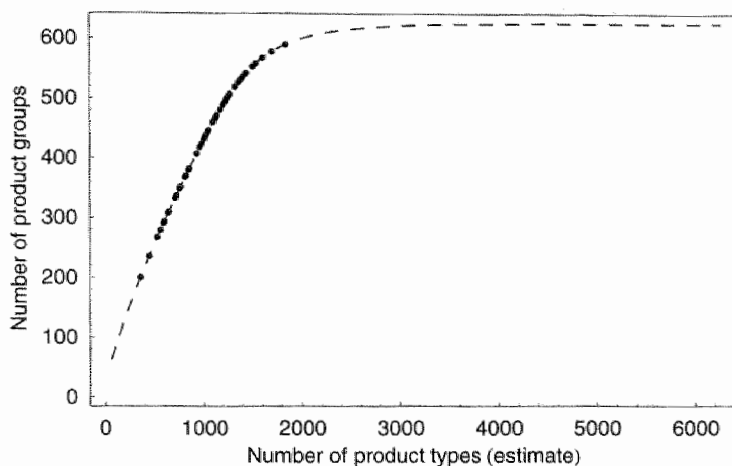
Standard deviations below coefficients

The results of this ad hoc experiment are displayed in figure 10.7. The clustering of types and groups causes the nearly piecewise-linear curve of figure 10.6 to become the smooth concave curve of figure 10.7.<sup>11</sup> Thus, the presence of clustering has a profound influence on the relation between the number of filled product groups and the number of product types.

It remains to be verified whether the extreme clustering illustrated by figure 10.7 is sufficient to explain the concavity in figure 10.2. The number of types estimated under the assumption of extreme clustering is plotted against population size in figure 10.8. For reference, also the estimated number of types without clustering has been plotted. Figure 10.8 demonstrates that even when clustering is pervasive, the relation between product variety and population size remains concave. However, accounting for clustering does reduce the concavity of the association between the number of types and population size, as can be seen from the estimation results presented in table 10.2. The coefficient on the log of population size is closer to one if clustering is assumed.

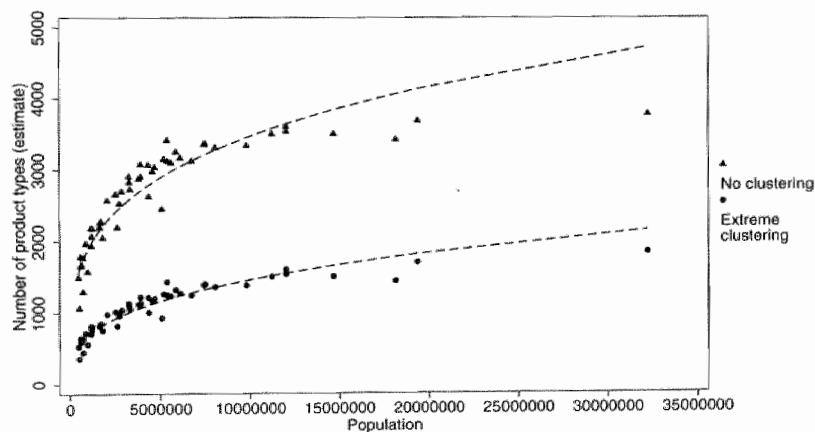
Summarizing, there is a fundamental difference between the observed number of filled product groups and the actual number of product types. This difference is essentially caused by the fact that the maximum number of product groups in a classification system is arbitrarily fixed. However, assuming a linear relationship between the number of groups and the number of types does not seem to be very inaccurate as long as the observed number of filled groups does not come close to the maximum number of groups.

The bias associated with the grouping of product types is unlikely to be the sole reason for the concave relation between product variety and population size that is observed in the cross-section of US states. The main conclusion of the previous section – that product variety in smaller states tends to be constrained by the optimal scale of production, while variety in large states is limited by the number of inventions – does not appear to be sensitive to the effects of product classification.



**Figure 10.7.** Estimated number of product types compared to actual number of product groups, assuming extreme clustering of groups and types (US states, 1997)

Sources: US Census Bureau; author's calculations



**Figure 10.8.** Estimated number of product types with and without clustering, and population size (US states, 1997)

Sources: US Census Bureau; author's calculations

## 10.4 Longitudinal determinants of product variety

A central prediction of the majority of the models discussed in the foregoing chapters is that the number of product types that is produced in an economy is linearly related to the size of the population – at least in the steady state. This dependence can simply be caused by the presence of fixed costs as in the original Dixit-Stiglitz framework (section 2.3). Alternatively, the number of types per capita can be constrained because products have to be invented like in the model by Judd (1985b). I will refer to models with a constant variety per capita as DSJ models.

The prediction of a constant number of types per capita does not apply in at least three cases. First, workers may become more productive when they use more machinery or are better educated like in the model of section 7.2. In this case product variety per capita is no longer constant as population size is no longer equivalent to the scale of the economy. A second case is constituted by growth models with intertemporal knowledge spillovers (chapter 4). In this class of models researchers benefit from previous research outcomes, which makes them invent new products faster. As a consequence, researchers in large autarkic economies are more productive than researchers in small economies.

A third mechanism that may prevent the number of types per capita to be constant is an ongoing change in schooling inequality. The model in section 7.2 demonstrates that a mean-preserving change in the distribution of skills will affect the amount of human capital devoted to research. Consequently, a changing skill distribution will also change the rate at which new product types are being introduced to the market.

In the remainder of this section, two tests are performed that should shed some light on which of the theories sketched above is the most appropriate. The first test evaluates whether the ratio of product variety to population size is constant; the second test evaluates whether the ratio of product variety to human capital is constant.

The cross-section analysis of section 10.2 has shown that the way in which countries interact is important for the number of product types they produce. Although these interaction effects are by themselves interesting, they also tend to obscure more fundamental – i.e. world-wide – determinants of product variety. In order to limit the influence of these interaction effects, this section will focus only on changes within countries. If the country-specific trends display a systematic pattern, then we might be able to draw some conclusions on what determines the world-wide set of products.

Unfortunately, if we want to do a time-series analysis of product variety, we encounter a new problem of data availability. The production data used for the cross-sectional analysis of section 10.2 is only available for a very small number of years. It is tempting to use trade data as a substitute for production data as detailed data on imports and exports have been collected consistently for at least the last forty years – but before doing so, we must take a closer look at the appropriateness of making such an approximation.

We can verify the reliability of trade data as a substitute for production data by regressing the number of produced groups,  $g_Y$ , on the ‘number of traded goods’ and a constant. The quality of the approximation can be judged by three criteria. First, the constant should be close to zero. Second, the standard deviation of the coefficient on the number of traded goods should be small. Third, the proportion of the variance in the number of produced goods that can be explained by the number of traded goods should be high.

**Table 10.3.** Approximation of the number of produced product groups (dependent variable) with the number of traded product groups

Sample $g_Y <$	Proxy	constant	coeff.	$R^2$			obs., $\rho$
				within	between	overall	
(full)	$g_X$	3.164	0.515	0.072	0.491	0.440	96
		354.363	0.107**				0.450
	$g_M$	131.605	0.427	0.043	0.359	0.319	96
		399.421	0.110**				0.459
	$g_{XM}$	139.818	0.409	0.042	0.350	0.312	96
		403.342	0.107**				0.458
3000	$g_X$	249.493	0.398	0.013	0.455	0.377	83
		332.705	0.105**				0.460
	$g_M$	394.286	0.314	0.003	0.331	0.287	83
		373.282	0.107**				0.461
	$g_{XM}$	397.261	0.301	0.003	0.327	0.290	83
		374.813	0.103**				0.460
2000	$g_X$	338.792	0.198	0.138	0.396	0.416	57
		168.528*	0.052**				0.720
	$g_M$	410.990	0.148	0.075	0.349	0.435	57
		176.474*	0.048**				0.719
	$g_{XM}$	393.813	0.148	0.079	0.344	0.439	57
		177.766*	0.047**				0.722

\*\* (\*) Coefficient significant at the 1% (5%) confidence level

Standard deviations below coefficients

Baltagi-Wu random effects estimation using EUROPROMS data

The 'number of traded goods' has been stated in quotation marks because trade data can be used in several ways. There are three obvious candidates that can function as a substitute for the number of produced groups. A first candidate is the number of exported groups,  $g_X$ . The number of exported groups is an appealing candidate as many of the goods that are produced in a country are also exported. Second, the number of imported groups,  $g_M$ , is attractive because many goods are produced with the use of imported intermediates.

A third candidate is the total number of groups that are exported, imported, or both,  $g_{XM}$ . Obviously, this candidate combines the advantages of the first two candidates. A disadvantage of this approach is that the number of groups is likely to get close to the maximum number of product groups for the larger countries, which makes this candidate sensitive to 'classification bias'.

Table 10.3 shows the results of regressing the number of produced groups on each of the three candidates separately. The underlying data stems from the EUROPROMS database. This database has the advantage that the data on imports and exports are classified in the same way as the production data. The estimation method is due to Baltagi and Wu (1999). This method corrects for autocorrelation in the residuals in the same way as the Prais-Winsten estimator ('partial differencing'), but then adapted to unbalanced panel data. The estimated autocorrelation coefficient,  $\rho$ , is reported in the last column of table 10.3.



The regressions have been repeated for three samples: the full sample, a medium-variety sample ( $g_Y < 3000$ ), and a low-variety sample ( $g_Y < 2000$ ). By comparing the approximation results for the three samples it can be verified whether approximation of production data with trade data is affected by classification bias.

Starting with the full sample, it is clear that  $g_X$  scores best on all of the three criteria mentioned earlier: it has the smallest constant, the highest t-value for  $g$ , and the highest  $R^2$ . For the medium-variety sample,  $g_X$  still scores best on all three points, but compared to the full sample the quality of approximation is lower. This also holds for the other two candidates. The results for the low-variety sample are mixed.  $g_X$  still has the smallest constant and the highest t-value, but the  $R^2$  of both  $g_M$  and  $g_{XM}$  is higher.

The results presented in table 10.3 suggest that the number of produced groups is best approximated by the number of exported groups. Restricting the sample does not yield a better approximation of produced variety: classification bias does not seem to affect the observed relation between produced and traded variety.<sup>12</sup>

Before testing the predictions of the various theories in detail, let us start by verifying whether product variety is actually growing. The longest time series of export data that is available, is classified according to SITC rev. 2. In order to verify whether the number of exported groups follows a random walk, a unit root test has been performed for each country. The breaks in the data caused by classification changes constitute a complication when performing these tests (see figures 3.12 and 3.13). Zivot and Andrews (1992) have developed a test that compares the null hypothesis of a unit root with the alternative hypothesis of a deterministic trend complemented by a single break in the series. The equation that has been used for detecting the presence of a unit root in the number of exported groups is stated below.<sup>13</sup>

$$g_{X,t} = \mu + \theta \text{SHIFT} + \beta t + \alpha g_{X,t-1} + \varepsilon_t \quad (10.5)$$

$$\text{SHIFT} \equiv \begin{cases} 0 & \text{if } t < \text{break} \\ 1 & \text{if } t \geq \text{break} \end{cases} \quad (10.6)$$

Here,  $t$  is time in years,  $\varepsilon_t$  is the residual, and SHIFT is a dummy variable that equals zero in the periods before the break and one from the period of the break onwards. The time at which the break occurs is determined endogenously by choosing the break that minimizes the t-statistic of  $\alpha$ . Of course, this procedure is only valid when there is at most one break in the series. For this reason, the sample has been restricted to the period 1970-1988, which is the longest period with only one break.

The results of the unit root tests are reported in the second and third column of table 10.4. The second column contains the t-value of  $\alpha$  that has been minimized by choosing the break period. The year in which the break occurs is listed in the third column. There are two perspectives on what are the appropriate critical t-values for  $\alpha$ . If the break is considered to be a random phenomenon, then the critical values are -5.34 for the 1% confidence level and -4.80 for the 5% level. If the break is considered to be an exogenous phenomenon, then the critical values are -4.32 and -3.76 (for a break halfway the sample; see table 2 of (Zivot and Andrews 1992)). The stars in table 10.4 are assigned using the latter critical values as changes in classification are not a direct reflection of changes in product variety. The null hypothesis of a unit root could not be refuted for only a handful of countries.<sup>14</sup>

**Table 10.4.** Unit root test and estimation of trend for the number of exported product groups reported by country for SITC 2 (1970-1988)

Country	$t_{\min}$	break	obs.	constant	SHIFT	$\hat{g}_X$
AUS	-5.882**	1979	18	-91.939 10.002**	0.289 0.045**	0.050 0.005**
AUT	-4.194*	1978	18	-69.629 9.154**	0.214 0.039**	0.039 0.005**
BEL	-15.945**	1978	18	-88.627 14.522**	0.343 0.057**	0.048 0.007**
CAN	-1.436	1973	18	-68.496 33.920	0.034 0.065	0.038 0.017*
CHE	-3.011	1978	18	-43.241 7.274**	0.135 0.023**	0.025 0.004**
DEU	-14.467**	1978	18	-86.989 14.667**	0.342 0.057**	0.048 0.007**
DNK	-18.254**	1978	18	-92.747 13.800**	0.345 0.056**	0.050 0.007**
ESP	-15.090**	1980	18	-100.335 13.309**	0.330 0.056**	0.054 0.007**
FIN	-18.643**	1978	18	-111.521 15.125**	0.389 0.064**	0.060 0.008**
FRA	-16.735**	1978	18	-86.687 14.472**	0.339 0.057**	0.047 0.007**
GBR	-18.307**	1978	18	-87.174 14.807**	0.343 0.059**	0.048 0.007**
GRC	-5.568**	1978	18	-143.623 15.145**	0.265 0.050**	0.076 0.008**
IRL	-3.538	1978	18	-206.854 50.120**	0.619 0.136**	0.108 0.025**
ISL	0.280	1975	17	-112.853 29.236**	-0.182 0.066*	0.059 0.015**
ITA	-15.363**	1978	18	-90.081 14.582**	0.347 0.057**	0.049 0.007**
JPN	-13.754**	1978	18	-88.501 14.929**	0.332 0.058**	0.048 0.008**
KOR	-4.126*	1973	12	-0.429 4.519		0.004 0.002
NLD	-20.631**	1978	18	-87.305 14.379**	0.346 0.058**	0.048 0.007**
NOR	-18.354**	1978	18	-106.751 14.129**	0.333 0.059**	0.057 0.007**
NZL	-14.479**	1979	18	-181.650 27.474**	0.583 0.115**	0.095 0.014**
PRT	-10.867**	1979	18	-95.371 12.568**	0.272 0.056**	0.052 0.006**
SWE	-14.609**	1978	18	-90.387 12.055**	0.332 0.056**	0.049 0.006**
TUR	-7.248**	1985	18	-162.035 35.194**	0.470 0.167**	0.085 0.018**
USA	-63.017**	1978	18	-70.264 10.046**	0.264 0.043**	0.039 0.005**

\*\* (\*) Coefficient significant at the 1% (5%) confidence level  
Standard deviations below coefficients

The unit root tests indicate that the growth rate of the number of exported groups can be estimated by fitting a deterministic trend to the data. The following equation has been used for the estimation of such trends.

$$\ln(g_{X,t}) = a_0 + a_1 \text{SHIFT} + a_2 t + \varepsilon_t \quad (10.7)$$

The method of estimation is OLS; the standard deviations of the coefficients are computed using Newey-West covariance estimates with two lags. The coefficient on the constant can be found in column 5, the coefficient on SHIFT in column 6, and the coefficient on time, which is an estimate of the growth rate of  $g_X$ , can be found in column 7. Standard deviations are reported below the coefficients.

The only two countries for which not all coefficients are significant are Canada and South Korea. Except for Iceland, all coefficients on SHIFT are positive and occur around 1978. This confirms that the breaks have been selected correctly for the other countries. More interesting, of course, are the estimates of the growth rate of the number of exported groups. This growth rate is positive for all countries as was predicted by theory. The first line of table 10.7 displays some summary statistics for table 10.4. The mean growth rate across countries of 5.3% per year appears to be reasonable and so does the standard deviation of 2.1% (also computed across countries).

Knowing that product variety has grown between 1970 and 1988 does not help us to discriminate between theories. As discussed at the beginning of this section, theories can be classified in two groups: those that predict a constant number of product types per capita and those that predict growth in this ratio. Which of the two groups has the highest empirical relevance can be established by estimating the trend in product variety per capita. Table 10.5 displays the results of unit root tests for the number of exported groups per capita as well as estimates of its growth rate,  $\hat{g}_X - \hat{P}$ . The estimation methods are the same as before.

The number of countries for which a unit root could not be refuted is larger than in table 10.4, but the number is still small enough to conclude that a unit root is absent in general. Like before the estimation results for Canada, Iceland, and South Korea are somewhat peculiar. In particular, only South Korea has a negative growth rate whereas this growth rate is positive for the other countries. The mean growth rate across countries is 4.2%; the standard deviation is 2.2% (table 10.7, line 2).

The clear upward trend in product variety per capita suggests that the models predicting a constant number of product types per capita are either too simple or simply wrong. In the beginning of this section, three theories were mentioned that can account for growth in variety per capita. The first proposes that the scale of an economy is determined by the stock of human capital rather than by the population size.<sup>15</sup> This augmented version of the DSJ model predicts a constant ratio of product variety to the stock of human capital.

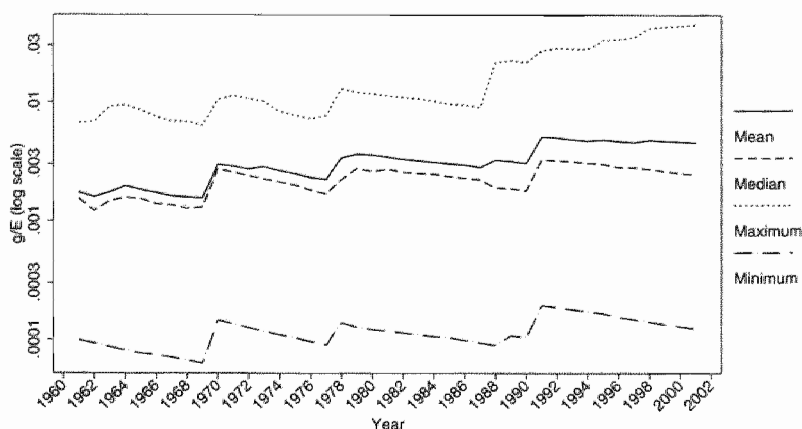
A second explanation is provided by models with intertemporal knowledge spillovers. These spillovers cause product variety per capita to grow through their effect on the productivity of researchers. If we take into account human capital, then, according to this class of models, also the ratio of product variety to the stock of human capital should be growing over time. The third explanation has to do with changes in schooling inequality. A more equal distribution of educational attainment may raise the number of people that decide to become a researcher.

We can evaluate the importance of intertemporal knowledge spillovers by conducting a second test. If the number of products per unit of human capital is constant, this can be seen as evidence in favor of the augmented DSJ model; if in-

**Table 10.5.** Unit root test and estimation of trend for the number of exported product groups per capita reported by country for SITC 2 (1970-1988)

Country	$t_{\min}$	break	obs.	constant	SHIFT	$\hat{g}_X - \hat{P}$
AUS	-6.994**	1979	18	-72.205 10.660**	0.286 0.047**	0.032 0.005**
AUT	-4.386**	1978	18	-70.332 9.138**	0.220 0.039**	0.031 0.005**
BEL	-17.711**	1978	18	-93.303 14.123**	0.342 0.056**	0.043 0.007**
CAN	-1.161	1974	18	-48.948 34.773	0.009 0.059	0.019 0.018
CHE	-3.491	1978	18	-47.040 7.084**	0.145 0.024**	0.020 0.004**
DEU	-15.368**	1978	18	-92.208 14.495**	0.347 0.057**	0.041 0.007**
DNK	-17.914**	1978	18	-98.756 13.734**	0.346 0.056**	0.046 0.007**
ESP	-14.104**	1980	18	-95.176 13.032**	0.331 0.055**	0.043 0.007**
FIN	-18.208**	1978	18	-116.156 14.966**	0.387 0.064**	0.055 0.008**
FRA	-17.831**	1978	18	-86.376 14.347**	0.342 0.057**	0.039 0.007**
GBR	-18.118**	1978	18	-96.122 14.630**	0.347 0.058**	0.043 0.007**
GRC	-5.708**	1978	18	-138.861 15.073**	0.268 0.049**	0.066 0.008**
IRL	-3.264	1978	18	-194.476 48.922**	0.611 0.133**	0.094 0.025**
ISL	-0.169	1975	17	-93.879 29.861**	-0.191 0.070*	0.044 0.015*
ITA	-15.715**	1978	18	-94.304 14.484**	0.354 0.057**	0.042 0.007**
JPN	-14.710**	1978	18	-89.474 15.094**	0.331 0.059**	0.040 0.008**
KOR	-3.702	1973	12	31.255 5.450**		-0.021 0.003**
NLD	-22.722**	1978	18	-78.448 13.902**	0.342 0.057**	0.035 0.007**
NOR	-19.228**	1978	18	-109.159 14.130**	0.333 0.059**	0.051 0.007**
NZL	-12.071**	1979	18	-168.596 28.243**	0.583 0.117**	0.081 0.014**
PRT	-10.951**	1979	18	-90.656 11.323**	0.263 0.053**	0.041 0.006**
SWE	-15.425**	1978	18	-101.747 12.007**	0.336 0.056**	0.047 0.006**
TUR	-6.639**	1985	18	-122.116 34.095**	0.455 0.161*	0.056 0.017**
USA	-50.238**	1978	18	-62.054 9.712**	0.251 0.042**	0.025 0.005**

\*\* (\*) Coefficient significant at the 1% (5%) confidence level  
Standard deviations below coefficients



**Figure 10.9.** Number of exported product groups per person with higher education completed (SITC 2/3, 1961-2001)

Sources: OECD (ITCS); UNSD (COMTRADE); Barro and Lee (2000); World Bank (WDI)

stead the ratio rises over time, then this will confirm the presence of intertemporal spillovers.

The stock of human capital is approximated here by the number of people that have completed higher education. Data on the proportion of the population that has completed higher education is available from the Barro-Lee data set (Barro and Lee 2000), but unfortunately only for every fifth year. In order to avoid a drastic reduction in the number of observations available for estimation, the data on education have been interpolated before being multiplied by the population size. The ratio of the number of product groups per highly educated person is plotted in figure 10.9. The discrete jumps are again due to classification changes. If anything, the trend in this ratio is slightly negative (taking into account changes in classification).

The outcomes of the unit root tests and the estimated growth rates can be found in table 10.6. The presence of a unit root can still be rejected for the majority of countries. There are only five countries for which the estimated trend in variety per highly educated person is significantly different from zero. For three of the countries the trend is negative; for two it is positive. The mean growth rate across countries is 0.2% – but with a standard deviation of 2.1% this is practically zero (table 10.7, line 3). The results presented in table 10.6 strongly suggest that the number of exported groups per highly educated person is constant. However, a closer look at the statistical significance of this outcome is useful.

The stars assigned in table 10.6 refer to the usual two-sided t-test of the null hypothesis that a coefficient is zero against the alternative hypothesis that it is not. Table 10.8 provides the p-values for a one sided t-test of the null hypothesis  $H_0 : \hat{g}_X - \hat{H} = 0$  against the alternative hypothesis  $H_1 : \hat{g}_X - \hat{H} > 0$ . For a confidence interval of 10% the alternative hypothesis can only be accepted for five countries – about a fifth of the sample. The mean p-value across countries is 46.3%, which is slightly lower than the 50% that could be expected when the growth rate would be zero on average.

**Table 10.6.** Unit root test and estimation of trend for the number of exported product groups per person with higher education completed reported by country for SITC 2 (1970-1988)

Country	$t_{\min}$	break	obs.	constant	SHIFT	$\hat{g}_X - \hat{H}$
AUS	-7.181**	1979	18	10.523 11.938	0.240 0.047**	-0.009 0.006
AUT	-5.403**	1978	18	12.530 9.129	0.181 0.036**	-0.009 0.005
BEL	-23.473**	1978	18	-10.441 13.398	0.303 0.054**	0.002 0.007
CAN	-0.107	1975	18	34.521 36.111	-0.042 0.059	-0.021 0.018
CHE	-2.983	1978	18	35.822 8.328**	0.106 0.024**	-0.021 0.004**
DEU	-23.990**	1978	18	-9.346 13.619	0.308 0.054**	0.001 0.007
DNK	-29.131**	1978	18	-15.894 12.692	0.307 0.053**	0.005 0.006
ESP	-22.882**	1980	18	-12.627 13.838	0.279 0.057**	0.003 0.007
FIN	-24.200**	1978	18	-33.295 13.857*	0.349 0.060**	0.014 0.007
FRA	-28.545**	1978	18	-3.514 13.520	0.303 0.054**	-0.002 0.007
GBR	-32.607**	1978	18	-13.260 13.636	0.308 0.055**	0.003 0.007
GRC	-6.426**	1978	18	-55.999 13.001**	0.229 0.043**	0.025 0.007**
IRL	-3.058	1973	18	-119.677 51.066*	0.483 0.231	0.058 0.026
ISL	-0.544	1974	17	-11.386 33.570	-0.139 0.101	0.003 0.017
ITA	-25.762**	1978	18	-11.442 13.469	0.316 0.054**	0.002 0.007
JPN	-18.327**	1978	18	-6.612 14.541	0.292 0.057**	-0.001 0.007
KOR	-4.395**	1972	12	101.138 8.201**		-0.055 0.004**
NLD	-29.114**	1978	18	4.414 13.219	0.303 0.054**	-0.005 0.007
NOR	-23.483**	1978	18	-26.298 13.382	0.295 0.056**	0.011 0.007
NZL	-12.250**	1979	18	-85.868 28.373**	0.537 0.116**	0.041 0.014*
PRT	-8.288**	1979	18	-7.928 13.001	0.217 0.056**	0.001 0.007
SWE	-18.984**	1978	18	-18.886 11.718	0.297 0.053**	0.007 0.006
TUR	-4.792**	1985	18	-38.771 37.905	0.485 0.179*	0.016 0.019
USA	-10.595**	1978	18	20.808 10.093	0.212 0.041**	-0.015 0.005**

\*\* (\*) Coefficient significant at the 1% (5%) confidence level  
Standard deviations below coefficients

**Table 10.7.** Summary of trends in the number of product groups for SITC 2, SITC 3, and HS 96

		mean	std dev.	min.	max.	obs.
SITC2 (1970-1988)	$\hat{g}_X$	0.053	0.021	0.004	0.108	24
	$\hat{g}_X - \hat{P}$	0.042	0.022	-0.021	0.094	24
	$\hat{g}_X - \hat{H}$	0.002	0.021	-0.055	0.058	24
	$\hat{g}_X - \hat{E}$	0.003	0.021	-0.053	0.059	24
SITC3 (1991-2001)	$\hat{g}_X$	0.004	0.021	-0.005	0.100	24
	$\hat{g}_X - \hat{P}$	-0.003	0.020	-0.022	0.089	24
	$\hat{g}_X - \hat{H}$	-0.034	0.020	-0.053	0.058	24
	$\hat{g}_X - \hat{E}$	-0.033	0.020	-0.052	0.057	24
HS96 (1996-2001)	$\hat{g}_X$	0.005	0.019	-0.005	0.102	29
	$\hat{g}_X - \hat{P}$	-0.003	0.020	-0.023	0.091	29

Summary statistics over countries for tables 10.4, 10.5, 10.6, 10.10, 10.11, and 10.12

**Table 10.8.** p-values t-test of  $\hat{g}_X - \hat{H} > 0$  (SITC 2)

Country	$p(T > t H_0)$	Country	$p(T > t H_0)$
AUS <sup>†</sup>	0.914	IRL	0.019
AUT <sup>†</sup>	0.969	ISL	0.422
BEL	0.364	ITA	0.385
CAN <sup>†</sup>	0.871	JPN <sup>†</sup>	0.544
CHE <sup>†</sup>	1.000	KOR <sup>†</sup>	1.000
DEU	0.455	NLD <sup>†</sup>	0.781
DNK	0.204	NOR	0.066
ESP	0.348	NZL	0.006
FIN	0.030	PRT	0.439
FRA <sup>†</sup>	0.610	SWE	0.139
GBR	0.338	TUR	0.215
GRC	0.001	USA <sup>†</sup>	0.995
mean	0.463	st. dev.	0.353

<sup>†</sup> negative growth rate  
 $H_0 : \hat{g}_X - \hat{H} = 0$  and  $H_1 : \hat{g}_X - \hat{H} > 0$

Altogether, the evidence is quite strong that the growth rate of product variety per highly educated worker is negligible. Of course, this conclusion hinges on a number of assumptions. First, export data are used in stead of production data. This implies that changes in trade policies and transport costs might lead to misestimation of variety growth. Second, the classification system and the time-interval might affect the results. Third, the measure of scale that has been used is still very crude.

In order to test the robustness of the findings, the exercise outlined above has been repeated for a second measure of human capital and for two other sets of export data. The alternative measure of human capital is the total number of years of higher schooling in a country,  $E$ . This measure also is based on data from Barro and Lee. Tables 10.7 and 10.10 (appendix 10.A) show that both measures of human capital lead to almost identical outcomes.

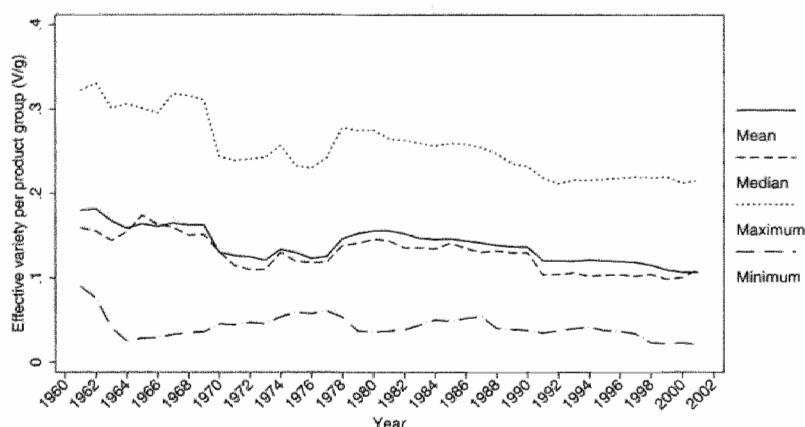
The first alternative set of data on exports covers the period from 1991 to 2001 and is classified according to SITC rev. 3. The estimation results for this data set are displayed in table 10.11. (See appendix 10.A for a description of the data). A summary of the table can be found in table 10.7. Remarkably, the growth rate of the number of exported groups is negligible. Given this observation, it is not surprising that the growth rates of variety per capita and variety per highly educated person are both negative. This is in contradiction with all of the theories considered above – only reduction in schooling inequality could provide a theoretical explanation for this observation.

The lack of growth in the number of product groups could be explained by a lack of coverage by the SITC classification system: many of the more innovative new products are not appropriately classified in SITC 3. This problem of coverage should be less severe for the Harmonized System (HS) classification. The outcomes for export data classified according to HS 96 can be found in tables 10.12 and 10.7 (see also appendix 10.A).<sup>16</sup> The estimation results for the HS 96 data are very similar to those obtained with the SITC 3 data. Apparently, the coverage of products is not a likely explanation for the constancy of the number of exported product groups.

The lack of growth in product variety observed in the 1990s contrasts starkly with the trends in the number of patents and research expenditure presented in chapter 3. It is therefore possible – or even likely – that the SITC 3 and HS 96 data sets do not give a reliable impression of the evolution of the actual number of product types in the countries under consideration. The question that remains to be settled is whether the results based on the SITC 2 data are more reliable than the results stemming from the other datasets.

If we are willing to accept the number of exported product groups in the SITC 2 dataset as an approximation of the actual number of produced types of goods, then a clear conclusion can be drawn about the empirical plausibility of the theories outlined above. The ordinary DSJ model is not plausible because product variety per capita has grown by roughly 4% per year on average. The DSJ model augmented with human capital, however, seems to be quite realistic as product variety per highly educated worker is constant for the majority of countries. The prediction of the model with intertemporal knowledge spillovers that variety per unit of human capital should grow over time does not seem to be realistic.





**Figure 10.10.** Effective variety per product group (exports, SITC 2/3, 1961-2001)

Sources: OECD (ITCS), UNSD (COMTRADE)

## 10.5 Idiosyncrasy and effective variety

Effective variety<sup>17</sup> plays a central role in two of the theoretical models presented earlier. The first was the model with imperfect improvements of chapter 7. Suppose that new product types are superior to older types, but also that they are not perfect substitutes for these older types. Although the superior new products attract a higher share of consumer spending than old products, old products continue to be bought as they provide some functionality that cannot be delivered by recent inventions. Consequentially, the number of product types will grow over time, provided that new types are being invented.

The superiority of new product types makes that consumers are no longer indifferent about the types they consume: if a person would be forced to abstain from using one type, he would prefer this type to be an old rather than a recent invention. The idiosyncrasy in the appreciation of types therefore warrants the effective variety of products to differ from the number of types. In the model of section 7.3 the variety index is indeed smaller than the number of products, as can be seen from equation 7.66. This equation also shows that the ratio of the variety index to the number of types does not change over time. Although every new invention might be better than the previous ones, this will not cause the effective variety per type to change.

Effective variety also plays a role in chapter 9. In this chapter the idiosyncrasy of product types arises because of random changes in the productivity of firms. Random productivity shocks cause the differences in the performance of firms widen over time. Naturally, consumers will buy more from relatively efficient firms as these firms will charge lower prices than their inefficient competitors. As in the 'imperfect improvements' case, consumers are not indifferent about which types they use, but now all types are of comparable quality and differ only with respect to their unit cost.

There is a second, and more important, respect in which the 'imperfect improvement' model and the 'idiosyncratic performance' model differ from one another. If one compares equations 7.66 and 9.38, then it will become clear that

whereas the ratio of effective variety to the number of types is constant in the first case, it is declining over time in the second. The presence of randomness has a fundamentally different effect on effective variety per type than the imperfect improvements have – at least within the context of the models.

Chapters 7 and 9 have demonstrated two ways by which a gap might arise between effective product variety and the number of types, but it remains to be seen whether such a gap is an important empirical phenomenon. The effective variety per product group,  $V/g$ , has been computed for each country and year using SITC export data.<sup>18</sup> Figure 10.10 summarizes the outcomes across countries.

A quick glance at the vertical scale reveals that the gap between effective variety and the number of products is substantial as the ratio  $V/g$  never exceeds 0.35. The influence of classification changes on the series plotted in figure 10.10 is not as pronounced as with the series on the number of product groups (fig. 3.12 and 3.13), but a clear break can still be discerned in the mean and median for the year 1991 – indicating the switch from SITC 2 to SITC 3.

Another remarkable feature of the series in 10.10 is that there is a lot of volatility before 1978, whereas afterwards the series behave much more regularly. For this reason the analysis in the remainder of this section will be confined to the years from 1978 onwards.

Table 10.9 reports a unit root test and an estimate of the growth rate of effective variety per product group. The table is constructed in the same way as table 10.4 of the previous section – only the variable of interest and the time interval are different. The *t*-values for the unit root test are clearly lower than those in tables 10.4, 10.5, and 10.6, but most of them are still high enough to reject the null hypothesis of a random walk.<sup>19</sup> The coefficient on the SHIFT-dummy is only significant for a few countries; when it is significant, the break tends to occur in 1988 or 1991.

The estimated growth rate of  $V/g$  reported in the last column is consistently negative, with the exception of New Zealand and Turkey. Of the negative trends, only two are not significant. The mean growth rate across countries is -2.0% per year; the standard deviation across countries is 2.1%. The mean might seem to be not substantially different from zero, but one has to take into account the skewness of the distribution: the 90% percentile of the distribution is still -0.2%. Altogether it seems safe to conclude that the gap between effective variety and the number of product groups is widening over time for almost all OECD countries. This finding is consistent with the model in chapter 9.<sup>20</sup>

The findings of this section do not provide any evidence on the plausibility of the 'imperfect improvement' model. This kind of evidence would be hard to obtain from the data that have been used. Of course, one could compare the average expenditure share of new product groups with that of old groups, but there are at least two reasons why these results will be biased. First, product specific R&D is likely to occur, such that new products arguably are not as close to their maximum functionality as older products are. Second, even when a new product is immediately of a higher quality than existing products, it will still take some time before the product will be sold on a large scale as buyers, sellers, and producers will have to adjust their habits to the new situation.

**Table 10.9.** Unit root test and estimated trend in effective variety per product group reported by country for SITC 2/3 (1978-2001)

Country	$t_{\min}$	break	obs.	constant	SHIFT	$\hat{V}_X - \hat{g}_X$
AUS	-6.042**	1988	23	50.443 3.540**	-0.156 0.022**	-0.027 0.002**
AUT	-5.002**	1988	23	7.390 4.660	0.118 0.015**	-0.005 0.002
BEL	-4.511**	1991	23	8.860 3.434*	-0.173 0.020**	-0.005 0.002**
CAN	-3.526	1996	23	75.269 14.901**	0.070 0.086	-0.039 0.008**
CHE	-3.143	1983	22	24.379 5.358**	0.036 0.026	-0.013 0.003**
DEU	-6.234**	1988	23	29.022 4.267**	0.063 0.015**	-0.015 0.002**
DNK	-4.586**	1991	21	7.611 2.758*	-0.053 0.014**	-0.005 0.001**
ESP	-6.468**	1990	22	37.212 8.856**	0.059 0.035	-0.020 0.004**
FIN	-4.815**	1985	21	38.049 11.560**	0.107 0.046*	-0.020 0.006**
FRA	-3.475	1988	23	24.134 4.388**	0.086 0.013**	-0.013 0.002**
GBR	-2.792	1997	23	48.369 7.288**	0.075 0.037	-0.025 0.004**
GRC	-3.729*	1987	23	19.576 7.849*	-0.007 0.035	-0.011 0.004*
IRL	-4.180*	1988	21	95.525 11.493**	0.146 0.031**	-0.049 0.006**
ISL	-10.226**	1988	23	186.192 11.868**	-0.357 0.073**	-0.095 0.006**
ITA	-3.487	1989	23	1.947 2.669	0.012 0.011	-0.002 0.001
JPN	-5.719**	1984	22	26.414 5.068**	-0.039 0.028	-0.014 0.003**
KOR	-4.625**	1986	23	57.012 8.398**	0.104 0.047*	-0.030 0.004**
NLD	-3.200	1991	23	46.450 4.862**	-0.080 0.018**	-0.024 0.002**
NOR	-7.005**	1988	23	51.328 11.066**	-0.101 0.043*	-0.027 0.006**
NZL	-5.792**	1984	23	-35.740 10.145**	0.113 0.056	0.016 0.005**
PRT	-2.571	1987	22	49.035 9.772**	0.035 0.039	-0.026 0.005**
SWE	-4.150*	1991	23	41.000 5.435**	-0.060 0.022*	-0.022 0.003**
TUR	-5.206**	1990	23	-3.550 11.589	0.049 0.045	0.001 0.006
USA	-4.180*	1983	23	20.765 6.176**	-0.010 0.040	-0.011 0.003**

\*\* (\*) Coefficient significant at the 1% (5%) confidence level  
Standard deviations below coefficients

## 10.6 Conclusion

The beginning of this chapter showed that the number of product groups for which a country's production is positive, is concavely related to the size of its population. This relation is particularly clear for the individual states of the US; the pattern for European countries is less pronounced. The only explanation for the concavity of this relation that is consistent with the theories discussed in this dissertation, is that product variety in small countries is constrained by the scale of production, while in large countries variety is constrained by the worldwide stock of inventions. This hypothesis is not easily refuted. Data on the number of establishments show that the concavity of the relation is not due to differences in the scale of production, and neither can the concavity be attributed entirely to classification effects.

Although the cross-section analysis indicates that the worldwide stock of inventions constrains the number of types in larger countries, the analysis does not provide any insights in what determines this stock of inventions. In order to get a grip on the more fundamental determinants of product variety, a time series analysis was performed. The time-series analysis, based on export data, shows that product variety per capita has grown over time in almost all of the OECD countries. The latter result appears to point towards the presence of intertemporal knowledge spillovers, but this conclusion turns out to be premature. Taking a measure of human capital as an indicator of the scale of an economy instead of population size, the support for intertemporal spillovers vanishes: the number of product groups per unit of human capital is constant over time for most countries.

Evidence presented at the end of the chapter shows that the ratio of effective product variety to the number of product groups declines over time. This finding can be explained theoretically by idiosyncratic firm performance. The impact of this phenomenon on consumer welfare, however, depends on the elasticity of substitution.

The following three conclusions can be drawn from these observations. First, the growth in the number of product types can be ascribed almost entirely to the growth of the stock of human capital, while intertemporal knowledge spillovers seem to be absent. Second, the evidence suggests that knowledge spillovers do occur between countries. Third, there is empirical support for the hypothesis that idiosyncratic shocks have a negative influence on the effective variety of products. This effect counteracts the positive effect on product variety attributable to the scale of the economy.

As a final remark, I would like to emphasize that the conclusions stated above are certainly not definitive. There are at least three reasons for concern. First, still little is known about how biased the number of product groups is as an indicator of the actual number of product types. Second, internationally comparable production data are only available for a couple of years, such that all longitudinal analyses have to rely on trade data. Third, all empirical studies completely ignore services. As services are responsible for a large part of GDP, caution must be taken in interpreting empirical trends.

10.A Additional figures and tables

Table 10.10 contains estimation results on the trend in product variety per unit of human capital and is similar to table 10.6, except for the fact that it uses an alternative measure of human capital: the total number of years of higher schooling that the inhabitants of a country have attained. Tables 10.11 and 10.12 contain estimates of  $\hat{g}_X$ ,  $\hat{g}_X - \hat{P}$ , and  $\hat{g}_X - \hat{H}$  (also reported in tables 10.4, 10.5, and 10.6), but then for SITC 3 and HS 96 export data.

Figures 10.11 and 10.12 contain scatterplots of the number of exported product groups plotted against population size for SITC 3 and HS 96, respectively. The figures show that the concave relation between variety and population size also holds for export data.

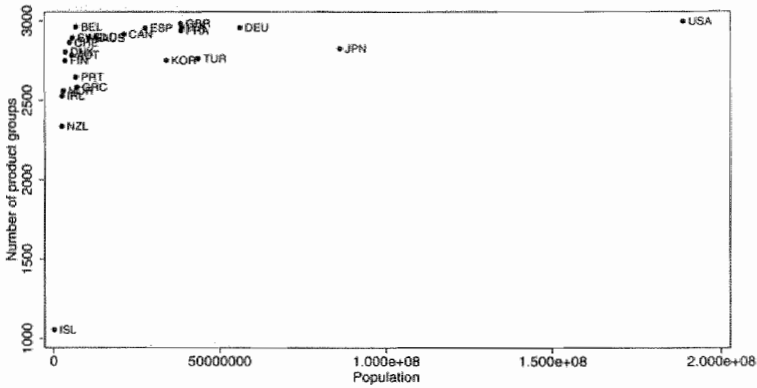


Figure 10.11. Number of product groups and population size for SITC 3 (exports, 2001)

Sources: OECD (ITCS); UNSD (COMTRADE); World Bank (WDI)

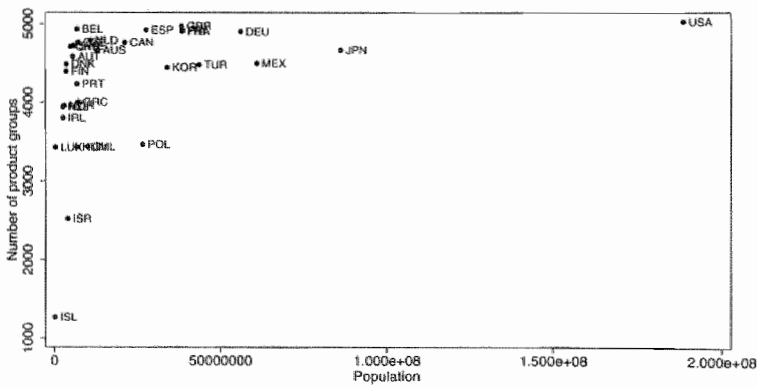


Figure 10.12. Number of product groups and population size for HS 96 (China excluded, exports, 2001)

Sources: OECD (ITCS); UNSD (COMTRADE); World Bank (WDI)

**Table 10.10.** Unit root test and estimation of trend for the ratio of the number of exported product groups to total years of schooling reported by country for SITC 2 (1970-1988)

Country	$t_{\min}$	break	obs.	constant	SHIFT	$\hat{g}_X - \hat{E}$
AUS	-6.974**	1979	18	7.698 12.625	0.236 0.049**	-0.008 0.006
AUT	-5.289**	1978	18	9.724 9.625	0.176 0.037**	-0.009 0.005
BEL	-21.102**	1978	18	-13.247 13.821	0.298 0.055**	0.003 0.007
CAN	-0.148	1975	18	31.845 36.395	-0.052 0.061	-0.021 0.018
CHE	-2.883	1978	18	33.016 8.871**	0.101 0.026**	-0.020 0.004**
DEU	-22.079**	1978	18	-12.153 14.007	0.303 0.055**	0.001 0.007
DNK	-28.528**	1978	18	-18.701 13.054	0.302 0.054**	0.006 0.007
ESP	-22.475**	1980	18	-15.461 14.441	0.277 0.058**	0.003 0.007
FIN	-26.288**	1978	18	-36.101 14.203*	0.344 0.061**	0.015 0.007
FRA	-25.330**	1978	18	-6.320 13.924	0.298 0.056**	-0.001 0.007
GBR	-31.253**	1978	18	-16.067 14.003	0.303 0.056**	0.003 0.007
GRC	-6.429**	1978	18	-58.806 13.049**	0.224 0.044**	0.026 0.007**
IRL	-3.022	1973	18	-122.468 50.273*	0.481 0.227	0.059 0.025*
ISL	-0.582	1974	17	-14.099 34.056	-0.146 0.104	0.004 0.017
ITA	-24.658**	1978	18	-14.249 13.826	0.310 0.055**	0.003 0.007
JPN	-16.552**	1978	18	-9.418 14.990	0.287 0.058**	-0.000 0.008
KOR	-4.408**	1972	12	96.439 7.983**	0.000 0.000	-0.053 0.004**
NLD	-24.974**	1978	18	1.608 13.659	0.298 0.056**	-0.005 0.007
NOR	-21.918**	1978	18	-29.104 13.797	0.290 0.058**	0.011 0.007
NZL	-12.546**	1979	18	-88.692 28.899**	0.533 0.117**	0.041 0.015*
PRT	-8.105**	1979	18	-10.753 13.706	0.213 0.057**	0.002 0.007
SWE	-18.239**	1978	18	-21.692 12.256	0.292 0.055**	0.007 0.006
TUR	-4.745**	1985	18	-41.502 38.467	0.489 0.180*	0.016 0.019
USA	-9.787**	1978	18	18.001 10.751	0.207 0.043**	-0.015 0.005*

\*\* (\*) Coefficient significant at the 1% (5%) confidence level  
Standard deviations below coefficients

Table 10.11. Estimation of trend for the number of exported product groups reported by country for SITC 3 (1991-2001)

Country	$g_X$		$g_X/P$		$g_X/H$	
	constant	$\hat{g}_X$	constant	$\hat{g}_X - \hat{P}$	constant	$\hat{g}_X - \hat{H}$
AUS	9.327 0.785**	-0.001 0.000	18.627 1.236**	-0.014 0.001**	83.564 1.655**	-0.045 0.001**
AUT	11.139 0.586**	-0.002 0.000**	4.420 1.054**	-0.006 0.001**	69.357 0.990**	-0.037 0.000**
BEL	11.375 0.550**	-0.002 0.000**	-0.930 0.642	-0.003 0.000**	64.007 1.009**	-0.035 0.001**
CAN	5.964 0.498**	0.001 0.000**	10.982 0.693**	-0.010 0.000**	75.918 0.870**	-0.041 0.000**
CHE	11.751 0.760**	-0.002 0.000**	5.551 1.087**	-0.006 0.001**	70.488 1.110**	-0.038 0.001**
DEU	13.978 0.753**	-0.003 0.000**	0.785 1.125	-0.005 0.001**	65.722 1.193**	-0.037 0.001**
DNK	15.792 2.870**	-0.004 0.001*	6.092 3.134	-0.007 0.002**	71.028 3.015**	-0.038 0.002**
ESP	10.265 0.970**	-0.001 0.000*	5.432 1.093**	-0.007 0.001**	70.369 1.491**	-0.039 0.001**
FIN	4.327 5.076	0.002 0.003	-4.298 5.068	-0.001 0.003	60.639 5.466**	-0.033 0.003**
FRA	17.453 1.161**	-0.005 0.001**	5.564 1.058**	-0.008 0.001**	70.501 0.988**	-0.039 0.000**
GBR	13.712 2.515**	-0.003 0.001*	0.654 2.560	-0.005 0.001**	65.590 2.909**	-0.036 0.001**
GRC	-16.951 4.877**	0.012 0.002**	-26.450 3.627**	0.009 0.002**	38.486 3.892**	-0.022 0.002**
IRL	16.844 8.137	-0.005 0.004	36.175 7.564**	-0.022 0.004**	101.111 7.226**	-0.053 0.004**
ISL	-192.700 9.908**	0.100 0.005**	-183.883 9.809**	0.089 0.005**	-118.946 9.433**	0.058 0.005**
ITA	13.960 0.840**	-0.003 0.000**	-4.417 1.015**	-0.003 0.001**	60.519 1.191**	-0.034 0.001**
JPN	12.575 0.589**	-0.002 0.000**	-5.770 1.276**	-0.002 0.001**	59.167 1.174**	-0.034 0.001**
KOR	3.105 2.219	0.002 0.001	11.201 1.399**	-0.010 0.001**	76.138 1.797**	-0.042 0.001**
NLD	15.694 1.104**	-0.004 0.001**	9.054 1.055**	-0.009 0.001**	73.990 1.339**	-0.040 0.001**
NOR	8.088 1.344**	-0.000 0.001	5.318 1.687*	-0.006 0.001**	70.254 2.055**	-0.038 0.001**
NZL	4.942 6.669	0.001 0.003	11.965 6.276**	-0.009 0.003*	76.902 6.456**	-0.041 0.003**
PRT	7.091 0.828*	0.000 0.000	-3.603 1.093**	-0.002 0.001**	61.333 1.015**	-0.034 0.001**
SWE	12.033 2.029*	-0.002 0.001	3.755 2.086	-0.006 0.001**	68.691 2.071**	-0.037 0.001**
TUR	-27.062 8.479**	0.018 0.004*	1.481 8.640	-0.006 0.004	66.417 8.948**	-0.037 0.004**
USA	12.632 0.712**	-0.002 0.000**	20.178 0.999**	-0.016 0.000**	85.114 1.402**	-0.047 0.001**

\*\* (\*) Coefficient significant at the 1% (5%) confidence level  
Standard deviations below coefficients

**Table 10.12.** Estimation of trend for the number of exported product groups reported by country for HS 96 (1996-2001)

Country	$g_X$		$g_X/P$	
	constant	$\hat{g}_X$	constant	$\hat{g}_X - \hat{P}$
AUS	2.790	0.003	14.512	-0.011
	0.617*	0.000**	0.693**	0.000**
AUT	14.579	-0.003	4.427	-0.006
	2.204**	0.001*	2.358	0.001*
BEL	7.053	0.001	-4.192	-0.002
	0.197**	0.000**	0.186**	0.000**
CAN	-0.611	0.005	5.670	-0.007
	4.472	0.002	4.496	0.002*
CHE	9.866	-0.001	0.551	-0.004
	0.635**	0.000	0.525	0.000**
CHL	1.316	0.003	16.910	-0.012
	5.295	0.003	5.175*	0.003*
CHN	7.782	0.000	8.000	-0.010
	0.766**	0.000	1.126**	0.001**
DEU	7.996	0.000	-7.242	-0.001
	0.332**	0.000	0.608**	0.000*
ESP	5.212	0.002	0.412	-0.005
	1.585*	0.001	1.529	0.001**
FIN	11.895	-0.002	3.076	-0.005
	1.695**	0.001	1.892	0.001**
FRA	11.563	-0.002	0.070	-0.005
	0.186**	0.000**	0.361	0.000**
GBR	5.247	0.002	-6.939	-0.001
	1.311*	0.001	1.210**	0.001
GRC	1.540	0.003	-12.385	0.002
	4.050	0.002	4.249*	0.002
HUN	-15.207	0.012	-34.717	0.014
	2.710**	0.001**	2.715**	0.001**
IRL	-12.382	0.010	8.506	-0.008
	3.308*	0.002**	2.612*	0.001**
ISL	-196.798	0.102	-187.301	0.091
	31.076**	0.016**	31.065**	0.016**
ISR	-13.796	0.011	26.275	-0.017
	29.119	0.015	28.584	0.014
ITA	11.571	-0.002	-8.464	-0.000
	0.339**	0.000**	0.717**	0.000
JPN	8.037	0.000	-13.862	0.002
	1.092**	0.001	0.812**	0.000**
KOR	4.881	0.002	9.272	-0.009
	0.630**	0.000**	0.769**	0.000**
MEX	17.772	-0.005	35.021	-0.022
	5.570*	0.003	6.542**	0.003**
NLD	16.954	-0.004	10.238	-0.009
	5.794*	0.003	5.841	0.003*
NOR	8.854	-0.000	7.444	-0.007
	1.793**	0.001	1.936*	0.001**
NZL	12.044	-0.002	17.496	-0.012
	2.976**	0.001	3.072**	0.002**

continued...



Table 10.12. (continued)

Country	$g_X$		$g_X/P$	
	constant	$\hat{g}_X$	constant	$\hat{g}_X - \hat{P}$
POL	-3.046	0.006	-5.493	-0.002
	1.429	0.001**	1.492*	0.001
PRT	8.172	0.000	-4.803	-0.001
	2.548*	0.001	2.669	0.001
SWE	0.352	0.004	-8.488	0.001
	2.196	0.001*	2.411*	0.001
TUR	8.030	0.000	36.693	-0.023
	2.635*	0.001	2.548**	0.001**
USA	7.876	0.000	17.829	-0.014
	0.150**	0.000*	0.342**	0.000**

\*\* (\*) Coefficient significant at the 1% (5%) confidence level  
Standard deviations below coefficients

## Notes

1. A 'product group' is a collection of, presumably similar, product types.
2. The seriousness of this last problem depends on the specific data set and estimation procedures that have been used.
3. It is assumed here that imitation is costless. This is the case when a firm produces the same good in different countries. When imitation does involve cost, population size will matter for produced variety, but this relation will be weak compared to the case of completely autarkic economies.
4. It is assumed that all economies are in a steady state and that the initial conditions are identical for all countries.
5. The apparent differences between Europe and the US are probably at least partly a consequence of the fact that the European data were not collected as consistently as the US data.
6. Figures 10.11 and 10.12 in appendix 10.A show that a concave relation between variety and size is also present in trade data.
7. Whether product  $a$  is the first or the last product that is assigned to group  $A$  is irrelevant. Neither does it matter whether group  $A$  and group  $B$  switch their positions or their names: the characteristics of a group depend purely on the types that are assigned to it.
8. If  $g = G$  then  $\lim_{n \rightarrow \infty} p(g|n, G) = 1$  such that the  $n$  that maximizes  $p$  is infinite.
9. A cubic spline did not lead to a substantially better fit.
10. None of the computations in this section take into account interactions between states, such as the geographical clustering of types. The number of types in a state is estimated using only the observed number of groups for that particular state and the maximum number of groups in the classification.
11. Formally both curves are strictly concave, but the slope of the curve in figure 10.7 declines more gradually than the slope in figure 10.6.
12. The numbers of exported and imported groups are much larger than the number of produced groups. This implies that classification bias is likely to be more severe in trade data than it is in production data.

13. Lagged differences have not been included because the corresponding coefficients were always insignificant.
14. Unit root tests on the log of  $g_X$  let to the rejection of a unit root for even more countries.
15. One might argue that mechanization – using more machines per worker – also affects the scale of the economy and that education is only part of the story. This is indeed likely, but from an empirical point of view it would be hard to separate the effects of education and mechanization because of the interaction between the two. The complementarity between mechanization and education, however, makes it unlikely that an adequate measure of human capital is a bad proxy of the scale of an economy – even when mechanization is important.
16. The HS 96 dataset covers only the period from 1996 to 2001, but for a larger sample of countries. The results related to human capital are omitted because the influence of the interpolation of schooling data would be very large.
17. Chapter 5 introduced the antilog of Shannon's entropy as a measure for the effective variety of products. This variety index differs from the 'number of product types' in that it accounts for the relative importance of each type; it can be considered a 'weighted number of types'. The variety index of chapter 5 can be used as an indicator of the contribution to welfare. It was shown in chapter 6 that a CES function can be decomposed into indices for variety, quality, and quantity.
18. The definition of the variety index  $V$  corresponds to the definition of  $N$  in equation 6.1, using the value of exports to compute the weights.
19. The stars assigned in table 10.9 are based on the critical values for an exogenous break, which depend on the timing of the break (Zivot and Andrews 1992, table 2). The t-values for an endogenous break are -5.34 (1%) and -4.80 (5%). Using these values it becomes much harder to reject a unit root: only nine countries have a t-value below -4.80.
20. By how much idiosyncratic firm performance would contribute to economic growth, depends on the elasticity of substitution. Combining equations 9.35 and 9.38, the part of the growth rate of consumption due to randomness is  $(1 - 2\gamma) / \gamma^2$  times  $\hat{V} - \hat{g}$ . This factor ranges between -1 for  $\gamma = 1$  and 0 for  $\gamma = 1/2$ , while it is positive for  $\gamma < 1/2$ . Assuming the model is correct, we still do not know anything if  $\gamma$  is not known.



## Conclusion

### 11.1 Summary

It has been the aim of this dissertation to clarify the relation between product variety and economic growth. A number of factors have been put forward that affect the nature of this relation. The most elementary of these factors, population size, was already discussed in chapter 2. In a nutshell, population size matters because a larger market will give producers a better opportunity to recover the expenditure required for the invention and development of new products. This opportunity becomes smaller as more products are introduced to the market. If the population does not grow, then the size of the market will be fixed and the number of product types will (eventually) be constant. The idea that market size and product variety are related is almost as old as the theory of monopolistic competition, but it was Judd (1985b) who first discussed the role of R&D.

A second factor has been introduced in chapter 4 in the form of intertemporal knowledge spillovers. If research outcomes improve the productivity of researchers, then a growing population is no longer essential for growth in product variety – provided that the spillovers from research to researchers are strong enough. Romer (1990) was the first to incorporate spillovers of this kind in a model of economic growth. As strong knowledge spillovers enable growth in product variety while population size is constant, they cause the number of product types per capita to grow. Jones (1995a) has demonstrated that when population growth is positive, also weak knowledge spillovers may lead to growth in product variety per capita. After Jones, a number of authors have proposed models with spillovers between quality-improving research and research aimed at the invention of new goods. Chapter 4 contains a generalized version of these models.

In chapter 7 two models of economic growth have been presented in which the traditional formulation of knowledge spillovers has been put aside. The first model is concerned with the effects of a third factor that affects product variety: education. The model shows that a rise in educational attainment leads to higher product variety per capita. Like population growth, advances in education raise the productive capacity of an economy thereby enabling the production of a greater number of goods. The model also shows that a change in schooling inequality affects product variety, although the direction of the effect is ambiguous. Schooling inequality is important for the amount of research being done because it affects the number of workers that prefer research to production work. A number of theoretical studies have been performed on the relation between education and

growth (Lucas 1988, Bils and Klenow 2000, Jones 2002), but none of them focus on the effects of schooling inequality.

A fourth determinant of variety is related to international trade. It is common in both the theoretical and empirical literature on economic growth to assume that each country has its own more or less independently functioning economy. This view has been criticized in chapter 3, where it is argued that especially advanced economies are highly integrated. A theoretical analysis of economic integration has been provided in chapter 8. One of the results presented here is that trade liberalization leads to a temporary increase in the growth rate in the number of products that is available world-wide. This effect occurs on top of the immediate rise in the variety of goods available for consumption due to the improved access to foreign markets – an effect emphasized by Romer (1994). The number of produced types in a country will decline as trade liberalization triggers international specialization and raises the scale of production.

In all of the models up to chapter 8, economic growth coincides with growth in product variety. Chapter 9 has demonstrated that this need not always be true. The models of this chapter show that permanent idiosyncratic shocks trigger a selection process, which results in a rising production for efficient firms and a declining production for inefficient firms. As a consequence, the aggregate quantity of goods increases while effective product variety decreases.

Clearly, there is a multitude of theories that can link product variety to economic growth. The propositions of chapter 1 make a distinction between two general mechanisms that affect both variety and the standard of living. The first mechanism comprises all factors that affect the scale of the economy – most notably population growth, advances in education, and economic integration. Proposition 1 states that these ‘scale factors’ have a positive influence on both product variety and the standard of living. This proposition has been tested empirically in sections 10.2 and 10.4. Cross-section analysis shows that for smaller countries population size is an important predictor of the produced variety of goods. The number of product types seems to be fairly constant across larger countries. This observation suggests that variety in small countries is constrained by a minimum efficient scale of production, while in large countries variety is constrained by the worldwide stock of inventions.

Although the cross-sectional evidence gives an indication that scale matters for variety, it does not provide a basis for deciding on the factors that determine the world-wide stock of inventions – for this a time-series analysis is required. Section 10.4 has shown that product variety per capita rises over time for most countries. Obviously, this means that the theories predicting a relation between variety and population size are not sufficient to explain the evolution of product variety. However, the ratio of the number of product types to a measure of human capital does remain constant over time. This observation contradicts the importance of intertemporal knowledge spillovers for product variety, while it emphasizes the role of education.

Putting the pieces of the puzzle together, both the cross-sectional and the time-series evidence are compatible with proposition 1. The cross-sectional pattern seems to be more accurately described by the model of chapter 8, while the results of the time-series analysis are more in line with the first model of chapter 7. Do these findings suggest a contradiction of theories? They do not. In fact, the models of chapters 7 and 8 are quite complementary: the analysis of economic integration of chapter 8 can also be performed while allowing for growth in educational attainment. Such an approach will probably not yield any new insights – it will just complicate the analysis.

Section 10.5 is concerned with testing proposition 2 of chapter 1, which states that idiosyncratic industry performance reduces effective product variety. The testing procedure builds upon a prediction of the model in continuous time of chapter 9. The prediction states that effective variety per product type should decline over time. The evidence presented in section 10.5 shows that effective variety per type is indeed declining over time for most countries. The second model of chapter 7 is also able to explain why effective variety can be smaller than the number of products, but it cannot account for growth in this difference.

Besides the theoretical and empirical support for propositions 1 and 2, this dissertation also contains a number of other results. For a start, the systematic analysis of intertemporal spillovers provided in chapter 4 has yielded a couple of models that appear to have been missing in the literature. Perhaps the most interesting of these two models demonstrates that constant growth in product variety is possible both with and without population growth. Of course, this outcome depends on very specific assumptions on the spillover structure.

The two chapters on measurement issues also contain some useful results. Chapter 5 has introduced an index of effective variety in the form of the antilog of Shannon's entropy. This index is superior to other indices because it can be disaggregated without losing its original functional form. Chapter 6 has presented a price and quantity index that are not biased by changes in product variety. The main properties of the indices do not rely on assumptions and they unconditionally pass the product test,  $E = PQ$ . The indices can be used to estimate changes in total factor productivity and the cost of living, provided that the production or utility function is homothetic. Within the domain of CES-functions, the proposed productivity and cost-of-living indices improve upon Feenstra's variety-index.

From a mathematical point of view, two – possibly new – theorems are of some interest: theorem 6 on multiplicative decomposition of sums, and theorem 10 on Jensen's inequality with a changing variance.

## 11.2 Further research

Naturally, further research has to be done on the plausibility of assumption 1: New goods tend to raise both the standard of living and the number of product types that are being produced. Despite the attention product variety receives in the literature on international trade, economic growth, industrial organization, and economic geography, very little is known by economists about how variety affects a consumer's preferences. Hardly any empirical work has been done and only two theoretical frameworks can cope with product variety: the Hotelling-Lancaster framework and the Dixit-Stiglitz framework. The first is cumbersome to use in a macroeconomic context and does not allow a consumer to consume more than one product type. The second framework is easy to use but lacks an intuitive foundation. Weitzman (1994) attempts to unify both approaches, but clearly a lot more needs to be done. Feenstra (2003) demonstrates that the translog expenditure function can also be used for modelling monopolistic competition. Empirical studies linking variety and consumer preferences are scarce. Bils and Klenow (2001) probably provide the only serious attempt.

When it comes to the relation between product variety and economic growth, the most serious challenges are of an empirical nature. More research needs to be done on the effects of product classification as this would help to give a better interpretation to econometric studies involving data on product variety. Classification bias can of course be avoided completely by using unclassified ('raw')

data. Detailed data on consumer products sold by retailers are being collected for commercial purposes, but the timespan covered by these data is rather short.

The analysis presented in this dissertation is primarily concerned with variety in consumer goods. The effects and determinants of variety in intermediate products are implicitly assumed to be governed by mechanisms that are similar to those of variety in consumer products. Obviously, additional research is required on the specificities of the two kinds of variety – both theoretically and empirically.

Another interesting opportunity for further research is to investigate more precisely what part of economic growth can be attributed to idiosyncratic shocks. Do economies grow because people wish to improve their situation, or is a large part of economic growth purely ‘accidental’?

A final remark is reserved for the measurement of economic growth. Traditionally, economic growth is measured by ‘real GDP’: the nominal value of a country’s production divided by a price index. ‘Real GDP’ is commonly interpreted as an indicator of aggregate welfare, but in reality it is just a quantity index. It is well-known that ‘real GDP’ captures only some indirect effects of changes in quality and variety. Nevertheless, ‘real GDP’ is still being used to test theories of economic growth as if it were a measure of welfare. If we want to take the theory seriously, we should not limit our attention to the quantity of goods but we should also take into account their quality and variety.

### 11.3 The Great Convergence

I end this concluding chapter with some thoughts on the future of the world-economy. However, before I will do so, I will make a brief detour to the past. One of the important questions asked by economic historians concerns the Industrial Revolution: Why did the Industrial Revolution take place in Britain? Why not somewhere else in Europe? Why not in China or India? Mokyr (1993) lists seven factors that made Britain a likely candidate for hosting the Industrial Revolution. In a nutshell, the most important of these factors are an abundance of coal, a relatively stable political climate, and a capability of refining and applying ‘macroinventions’ stemming from continental Europe.<sup>1</sup>

Although the Industrial Revolution took place in Britain, it certainly did not take place in isolation from the other European countries. At the beginning of the Industrial Revolution Britain exported around 15 percent of its GNP.<sup>2</sup> Not only goods were exchanged between the countries of Europe, also knowledge about inventions and scientific results easily crossed national borders. The spinning techniques that greatly improved the productivity of the textile sector in Britain were based on French inventions, while the agricultural techniques that freed up labor for industrial production were imported from The Netherlands.<sup>3</sup>

Landes (1998, chap. 14) argues that the climate fostering the accumulation, sharing, and application of knowledge present in Europe was lacking in China and India. The Chinese were not interested in foreign technologies, because they considered all things foreign to be inferior. Moreover, technological change was not in the interest of the Chinese emperor, whose celestial status could be undermined by earthly innovations. India, the other populous country in Asia, did not welcome new technologies either. The sheer abundance of labor in India gave nobody an incentive to use more efficient technologies. Kerr (1995, pp. 95-7) gives a striking example of Indian workers carrying wheelbarrows on their heads in order to provide employment for their women and children.<sup>4</sup>

In his book "The Great Divergence", Pomeranz (2000) emphasizes other factors that gave Britain an advantage over China. In particular, China lacked easy access to coal and, unlike Britain, did not have colonies that could be exploited. He argues that it was the availability of resources, rather than knowledge, that made the difference. Pomeranz points out that also the Chinese were knowledgeable about the uses of steam and coal, but that they did not need to build steam engines in order to keep their mines from flooding.

Mokyr (1990, chap. 9), like Landes, emphasizes the role of the state in China. The state's attitude to technology was all-important but not necessarily negative. The Chinese rate of technological change surpassed that of Europe during the Sung (960-1126) and Mogol (1127-1367) dynasties, but came to a standstill during the Ming (1368-1644) and Qing (1644-1911) dynasties. The risk of such a standstill of technology was much smaller in Europe than it was in China because Europe consisted of several competing states that were not powerful enough to prevent the private development and use of new technologies.

Whatever the exact reasons for the divergence in the standard of living between Europe and Asia, it is clear that this divergence is substantial and persistent. Even though more than two centuries have passed since the start of the Industrial Revolution, the economies of China and India are still largely isolated from the world-economy. I will not attempt to analyze in detail why the Great Divergence has been so persistent, but there are some likely causes. Communism in general and Mao Zedong in particular have been detrimental to economic growth in China, and India is still a very populous country with a rigid caste system and a traditional division of labor.

This thesis began with an overview of the early stages of what can be called the 'world-economy'; it will end with a discussion of its future. The single most important question regarding the future of the world-economy is for how long the Great Divergence will persist. Why is this question so important? The facts are clear. The size of the population of the 24 OECD countries combined is about 1 billion; the combined population of China and India is approximately 2.3 billion. Furthermore, the growth rate of the population is 0.7 percent for the OECD, while it is 1.1 percent for China and India together.<sup>5</sup> These numbers are sufficiently impressive to justify the conclusion that even a moderate liberalization of the Chinese and Indian economies would increase the scale of the world-economy substantially. According to proposition 1 and (especially) chapter 8 of this thesis, an increase in the scale of an economy will stimulate the invention of new goods and will therefore raise the standard of living of its inhabitants. An increase in the size of world-markets implies that it becomes more profitable to invent and develop new products as they can be sold to a greater number of customers. The integration of China and India into the world-economy will therefore not only bring about welfare gains for their own inhabitants, but it will also cause a rise in the growth rate of per capita income at the global level.

Two remarks are in place regarding the pace and impact of the integration of China and India into the world-economy. First, the incorporation of countries of this size will not happen overnight – if it happens at all. The economic systems of the two countries are quite different from those of the OECD-countries, so that it is likely that the process of integration would take several decades to complete. Economic liberalization will typically come gradually and will at first be limited to specific areas (China: Guangdong Province) or sectors (India: information technology).



Second, the average productivity of workers in China and India is considerably lower than in OECD countries. For this reason, population size is a biased indicator of the potential gain in scale that can be accomplished through integration. The differences in productivity stem from roughly two sources. First, the average level of schooling is lower. It will take several decades before educational attainment in China and India will be at the same level as in economically advanced countries. Second, labor is not employed as efficiently as in OECD countries because of, for example, improperly functioning markets, restrictive regulation of businesses, or social hierarchies like the caste system and, perhaps, the Communist Party.

What are the odds that a rapid expansion of the world-economy will occur? The principal answer to this question lies in the nature of the reforms that China and India have started. Will these reforms lead to full modernization of the economy, or will they merely prevent its deterioration? Let us start with China. Already a few years after the end of the Cultural Revolution (1966-1976) had ended, the Chinese government started to reform the economy. It did so because it was dissatisfied with the poor performance of the economy relative to that of some of China's neighboring countries, and not because it wanted to imitate European or other 'western' countries. The reforms of the early 1980s were aimed at decentralization of administrative and economic power. Firms were allowed to retain a fraction of their profits, and local governments gained greater autonomy over their budgets and economic policy.<sup>6</sup>

The decentralization of economic power did not coincide with the abolition of central fixation of prices. This strengthened the incentive for local governments to restrict interregional trade and triggered the duplication of industries. Decentralization thus led to fractionalization of the economy.<sup>7</sup> The duplication of industries came to a halt only halfway through the 1990s.

China's export of manufactures has increased substantially since 1980, although reliable numbers are missing (Perkins 1994, p. 33-34). This rise has occurred despite the fact that all trade with other countries was handled by foreign trade corporations, which isolated domestic companies from their foreign partners. From the beginning of the 1990s onwards, also some domestic firms were allowed to trade directly with foreign firms and to retain part of the foreign currency they earned.

The increasing autonomy of firms and local governments demonstrates that the Chinese government is committed to the reform of the economy. China's admission to the WTO in 2001 shows further proof that China is willing to participate in the world-economy. However, reform by the central government proceeds gradually and the protectionism of local governments implies that China will open up region by region, which will further slow down the process of integration.

The reform of the Indian economy started off in 1991 – much later than was the case for China.<sup>8</sup> The severe balance of payments crisis of that year forced the government to start with reforming the economy. Unlike China, India had the advantage that it did not inherit a communist economy: private property and markets were well-spread. This, however, did not prove to lead to more rapid reforms. The problem of the Indian economy was not the absence of markets but the extremely detailed and restrictive regulation of private investment. During the reforms of the 1990s the number of rules was reduced substantially. For example, industrial licensing by the central government was abolished for most industries, and only three industries remained reserved for the public sector (down from 17). However, the production of many goods is still allowed on a small scale only, and

with technologies prescribed by the government. Restrictions were lifted for some of these goods in 2000 and 2001, but remain in tact for many products.

In 1993, quantitative restrictions on the import of capital goods and intermediates were abolished, but the quantitative restrictions on the import of final products were not removed until 2001 – when they were replaced by tariffs. Although the share of exports in GDP has risen in the last 15 years, it still is a meager 14 percent.<sup>9</sup> Government policies designed to attract foreign investors did not stimulate exports like they did in China. There is one well-known exception to these rather disappointing results. India's economic reform has been successful in stimulating the export of services related to and enabled by information technology (e.g. call-centers and software development). Various factors underlie this success: absence of industry regulation, low duties on required intermediates, and an improved telecommunications infrastructure.

When compared with China, India's economic reform has been late and even more gradual. A likely reason for the slow pace of change is the fact that reforms are difficult to get approved in India's democracy – a nuisance that does not bother China. As a consequence, India's reforms have been defensive rather than offensive like in China. Judging by the experiences of the last twenty years it seems quite likely that China will continue the gradual liberalization of its economy, while India will stay reluctant to do so. Nevertheless, the sheer size of both countries will make up for the slow pace of their reforms. Even a gradual and partial integration of China and India will stimulate the invention of new goods for the next fifty to hundred years.

## Notes

1. Later, Mokyr (2002, p. 77) argues that the Industrial Revolution could also have happened in other European countries.
2. See Mokyr (1993, p. 68). Chapter 1 contains a historical overview of international trade in Europe.
3. See Landes (1998, p. 214) and Mokyr (1993, p. 21).
4. See also Landes (1998, p. 229).
5. The data for population size and population growth are for 2001 and stem from the World Development Indicators (World Bank 2003).
6. See Jefferson and Rawski (1994) and Young (2000).
7. Young (2000) discusses the consequences of decentralization in detail.
8. The next paragraphs draw heavily on Ahluwalia (2002).
9. Data for 2001 from the World Development Indicators (World Bank 2003).



## Data on product variety

The detailed production data used in this dissertation stem from two sources. First, production data for 51 states of the US are available from the 1997 Economic Census conducted by the US Census Bureau. The used data set contains 626 industries and can be downloaded from the internet ([www.census.gov](http://www.census.gov)). The second source of production data is Statistics Europe (2001), which is a joint venture of Eurostat and the statistical agencies of the 16 participating countries. Data are classified according the Prodcom nomenclature (about 5000 different economic activities), which is comparable to the Combined Nomenclature. The database also contains trade data at the same level of detail.

Two datasets have been compiled containing the value of imports and exports by product group, country, and year. The first dataset covers the 24 OECD member countries for the years 1961 through 2001 and is based on the SITC system for product classification. Section 3 of SITC ("Mineral fuels, lubricants and related materials", 67 product groups) has been excluded because it has a large impact on the entropy of trade values. The SITC classification has been revised three times. Officially, the data for the period 1961 to 1990 are classified according to SITC revision 2, while the data after 1990 are classified according to revision 3. Unofficially, the data stemming from before 1976 originate from revision 1. This can be concluded from figure 3.14 and has been confirmed by the OECD.

The database has been compiled from two sources. The primary source for this dataset is the OECD's International Trade by Commodity Statistics (ITCS) database as available from CD-ROM (OECD 2002a, OECD 2002b). For some combinations of countries and years data was missing, therefore also data has been used from the United Nations Statistics Division's (UNSD) COMTRADE database (United Nations Statistics Division 2004). Details on the compiled database can be found in table A.2. A list explaining the codification of the countries is provided in table A.1.

The second dataset covers 32 countries from 1996 to 2001. For this dataset, products are classified according to the 1996 version Harmonizes System (HS 96), which is more detailed than the SITC classification (6 digits in stead of 5 digits). Chapter 27 of HS ("Mineral fuels, mineral oils and products of their distillation; bituminous substances; mineral waxes", 58 product groups) has been excluded. This dataset has also been compiled from the ITCS and COMTRADE databases. Details can be found in table A.3.

**Table A.1.** ISOALPHA 3 country codes

AUS	Australia	IRL	Ireland
AUT	Austria	ISL	Iceland
BEL	Belgium	ISR	Israel
CAN	Canada	ITA	Italy
CHE	Switzerland	JPN	Japan
CHL	Chile	KOR	Republic of Korea
CHN	China	LUX	Luxembourg
CZE	Czech Republic	MEX	Mexico
DEU	Germany	NLD	Netherlands
DNK	Denmark	NOR	Norway
ESP	Spain	NZL	New Zealand
FIN	Finland	POL	Poland
FRA	France	PRT	Portugal
GBR	United Kingdom	SWE	Sweden
GRC	Greece	TUR	Turkey
HUN	Hungary	USA	United States

Table A.2. DATA COVERED SITC rev. 2 and 3

	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981
AUS			2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
AUT	2																				2
BEL		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1
CAN	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
CHE	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
DEU	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
DNK	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
ESP	2		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
FIN																					2
FRA	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
GBR	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
GRC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
IRL	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
ISL	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
ITA	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
JPN		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
KOR																1	1	1	1	1	1
NLD	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
NOR	2		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
NZL																					2
PRT	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
SWE	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
TUR	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
USA	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
AUS	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
AUT	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BEL	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
CAN	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1
CHE	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1
DEU	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
DNK	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
ESP	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
FIN	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
FRA	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
GBR	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
GRC	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
IRL	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1
ISL	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1
ITA	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
JPN	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
KOR	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
NLD	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
NOR	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
NZL	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
PRT	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
SWE	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
TUR	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
USA	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

SITC rev. 2 from 1961 to 1990; SITC rev. 3 from 1991 to 2001  
Sources: 1 UNSD, 2 OECD

Table A.3. Data coverage HS 96

	1996	1997	1998	1999	2000	2001
AUS	2	2	2	2	2	2
AUT	1	1	1	1	1	1
BEL	2	2	2	2	2	2
CAN	2	2	2	2	2	2
CHE	2	2	2	2	2	2
CHL		1	1	1	1	1
CHN	2	2	2	2	2	2
CZE	1	1*	1	1	1	1
DEU	1	1	1	1	1	1
DNK	1	1*	1	1	1	1
ESP	1	1	1	1	1	1
FIN	2	2	2	2	2	2
FRA	2	2	2	2	2	2
GBR	2	2	2	2	2	2
GRC	1	1	1	1	1	1
HUN	1	1	1	1	1	1
IRL	2	2	2	2	2	2
ISL	2	2	2	2	2	2
ISR	1	1	1	1	1	1
ITA	2	2	2	2	2	2
JPN	2	2	2	2	2	2
KOR	2	2	2	2	2	2
LUX				1	1	1
MEX	1	1	1	1	1	1
NLD	1	1	1	1	1	1
NOR	2	2	2	2	2	2
NZL	1	1	1	1	1	1
POL	2	2	2	2	2	2
PRT	1	1	1	1	1	1
SWE	1	1	1	1	1	1
TUR	1	1	1	1	1	1
USA	2	2	2	2	2	2

Sources: 1 UNSD, 2 OECD

\* Export data not available

# Bibliography

- Abramowitz, M., and I. A. Stegun (eds.) (1972): *Handbook of mathematical functions*. Dover Publications, inc., New York.
- Acemoglu, D. (1998): "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *Quarterly Journal of Economics*, 113(4), 1055–89.
- (2002): "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 40(1), 7–72.
- Aghion, P., and P. Howitt (1992): "A Model of Growth through Creative Destruction," *Econometrica*, 60(2), 323–51.
- (1998): *Endogenous growth theory*. MIT Press, Cambridge and London.
- Ahluwalia, M. S. (2002): "Economic Reforms in India since 1991: Has Gradualism Worked?," *Journal of Economic Perspectives*, 16(3), 67–88.
- Alexander, P. J. (1997): "Product Variety and Market Structure: A New Measure and a Simple Test," *Journal of Economic Behavior and Organization*, 32(2), 207–14.
- Amoroso (1930): "La curva statica di offerta," *Giornale degli Economisti*.
- Arnold, L. G. (1998): "Growth, Welfare, and Trade in an Integrated Model of Human-Capital Accumulation and Research," *Journal of Macroeconomics*, 20(1), 81–105.
- Balk, B. M. (1998): *Industrial price, quantity, and productivity indices: The microeconomic theory and an application*. Kluwer Academic, Boston; Dordrecht and London.
- (2000): *On curing the CPI's substitution and new goods bias*. Department of Statistical Methods Research Paper 0005, Statistics Netherlands, Voorburg.
- Baltagi, B. H., and P. X. Wu (1999): "Unequally Spaced Panel Data Regressions with AR(1) Disturbances," *Econometric Theory*, 15(6), 814–23.
- Barro, R. J., and J. W. Lee (2000): "International Data on Educational Attainment Updates and Implications," Discussion Paper 7911, National Bureau of Economic Research.
- (2001): "International Data on Educational Attainment: Updates and Implications," *Oxford Economic Papers*, 53(3), 541–63.
- Barro, R. J., N. G. Mankiw, and X. Sala-i Martin (1995): "Capital Mobility in Neoclassical Models of Growth," *American Economic Review*, 85(1), 103–15.
- Barro, R. J., and X. Sala-i-Martin (2004): *Economic growth*. MIT Press, Cambridge (MA) and London, 2 edn.
- Becker, G. S. (1965): "A theory of the allocation of time," *Economic Journal*, 75(299), 493–517.



- Ben-David, D. (1993): "Equalizing Exchange: Trade Liberalization and Income Convergence," *Quarterly Journal of Economics*, 108(3), 653–79.
- Berry, S., J. Levinsohn, and A. Pakes (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112(1), 68–105.
- Bils, M., and P. J. Klenow (2000): "Does Schooling Cause Growth?," *American Economic Review*, 90(5), 1160–83.
- (2001): "The Acceleration of Variety Growth," *American Economic Review*, 91(2), 274–80.
- Bowley, A. L. (1899): "Wages, nominal and real," pp. 640–651. Macmillan, London.
- Brakman, S., and H. Garretsen (2003): "Rethinking the 'New' Geographical Economics," *Regional Studies*, 37(6–7), 637–48.
- Bresnahan, T. F., and M. Trajtenberg (1995): "General Purpose Technologies: 'Engines of Growth'?", *Journal of Econometrics*, 65(1), 83–108.
- Broda, C., and D. E. Weinstein (2004a): *Globalization and the Gains from Variety*. NBER Working Paper 10314.
- (2004b): "Variety Growth and World Welfare," *American Economic Review*, 94(2), 139–44.
- Brunner, M., and H. Strulik (2002): "Solution of Perfect Foresight Saddlepoint Problems: A Simple Method and Applications," *Journal of Economic Dynamics and Control*, 26(5), 737–53.
- Brynjolfsson, E., and L. M. Hitt (2003): "Computing Productivity: Firm-Level Evidence," *Review of Economics and Statistics*, 85(4), 793–808.
- Calvo-Armengol, A., and Y. Zenou (2002): "The Importance of the Distribution of Consumers in Horizontal Product Differentiation," *Journal of Regional Science*, 42(4), 793–803.
- Cass, D. (1965): "Optimum Growth in an Aggregative Model of Capital Accumulation," *Review of Economic Studies*, 32(July), 233–240.
- Chamberlin, E. H. (1937): "Monopolistic or imperfect competition?," *Quarterly Journal of Economics*, 51(4), 557–580.
- (1950): "Product heterogeneity and public policy," *American Economic Review*, 40(2), 85–92.
- (1962 [1933]): *The theory of monopolistic competition*. Harvard University Press, Cambridge, 8th edn.
- Conlisk, J. (1989): "An Aggregate Model of Technical Change," *Quarterly Journal of Economics*, 104(4), 787–821.
- Crafts, N. F. R., and C. K. Harley (1992): "Output Growth and the British Industrial Revolution: A Restatement of the Crafts-Harley View," *Economic History Review*, 45(4), 703–30.
- Davidson, J. (1994): *Stochastic limit theory: An introduction for econometricians*. Oxford University Press, Oxford and New York.
- Devereux, M. B., and B. J. Lapham (1994): "The Stability of Economic Integration and Endogenous Growth," *Quarterly Journal of Economics*, 109(1), 299–305.
- Diewert, W. E. (1980): "Aggregation problems in the measurement of capital," in *The measurement of capital*, ed. by D. Usher. University of Chicago Press, Chicago and London.

- (1993): "Overview of Volume I," in *Essays in Index Number Theory Volume 1*, ed. by W. E. Diewert, and A. O. Nakamura. North-Holland, Amsterdam.
- (2001): "The Consumer Price Index and Index Number Theory: A Survey," Dep. of Ec. Discussion Paper 01/02, University of British Columbia.
- Dinopoulos, E., and P. Thompson (1998): "Schumpeterian Growth without Scale Effects," *Journal of Economic Growth*, 3(4), 313–35.
- Dixit, A. K. (2000): "Some reflections on theories and applications of monopolistic competition," in *SOM conference on 'The monopolistic competition revolution after twenty-five years'*, Groningen. University of Groningen.
- Dixit, A. K., and V. Norman (1980): *Theory of international trade*. Cambridge University Press, Cambridge.
- Dixit, A. K., and J. E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67(3), 297–308.
- (1993): "Monopolistic Competition and Optimum Product Diversity: Reply," *American Economic Review*, 83(1), 302–04.
- Drobisch (1871): "Ueber Mittelgrößen und die Anwendbarkeit derselben auf die Berechnung des Steigens und Sinkens des Geldwerthes," *Berichte der mathematisch-physikalischen Classe der Königlich Sächsischen Gesellschaft der Wissenschaften*.
- Duesenberry, J. S. (1949): *Income, saving and the theory of consumer behavior*. Cambridge University Press, London.
- Dupuy, A., and P. Marey (2004): *Shifts and Twists in the Relative Productivity of Skilled Labor: Reconciling Accelerated SBTC with the productivity Slowdown*. Unpublished manuscript.
- Eichhorn, W. (1978): *Functional Equations in Economics*. Addison-Wesley Publishing Company, Reading, MA.
- Estevadeordal, A., B. Frantz, and A. M. Taylor (2003): "The Rise and Fall of World Trade, 1870-1939," *Quarterly Journal of Economics*, 118(2), 359–407.
- Ethier, W. J. (1982): "National and International Returns to Scale in the Modern Theory of International Trade," *American Economic Review*, 72(3), 389–405.
- Feenstra, R. C. (1994): "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 84(1), 157–77.
- (2003): "A Homothetic Utility Function for Monopolistic Competition Models, Without Constant Price Elasticity," *Economics Letters*, 78(1), 79–86.
- Feenstra, R. C., and H. L. Kee (2004): "On the Measurement of Product Variety in Trade," *American Economic Review*, 94(2), 145–49.
- Feenstra, R. C., D. Madani, T.-H. Yang, and C.-Y. Liang (1999): "Testing Endogenous Growth in South Korea and Taiwan," *Journal of Development Economics*, 60(2), 317–41.
- Feenstra, R. C., and J. R. Markusen (1994): "Accounting for Growth with New Inputs," *International Economic Review*, 35(2), 429–47.
- Feenstra, R. C., J. R. Markusen, and W. Zeile (1992): "Accounting for Growth with New Inputs: Theory and Evidence," *American Economic Review*, 82(2), 415–21.
- Feenstra, R. C., T. H. Yang, and G. G. Hamilton (1999): "Business Groups and Product Variety in Trade: Evidence from South Korea, Taiwan and Japan," *Journal of International Economics*, 48(1), 71–100.

- Fisher, F., and K. Shell (1972): *The economic theory of price indices*. Academic Press, New York and London.
- Fisher, R. A. (1958 [1930]): *The genetical theory of natural selection*. Dover Publications, New York, 2d rev. edn.
- Flam, H., and E. Helpman (1987): "Vertical Product Differentiation and North-South Trade," *American Economic Review*, 77(5), 810–22.
- Flynn, D. O. (2001): "The Great Divergence: China, Europe, and the making of the modern world economy (review)," *Journal of Economic Literature*, 39(2), 606–08.
- Frenken, K., P. P. Saviotti, and M. Trommetter (1999): "Variety and Niche Creation in Aircraft, Helicopters, Motorcycles and Microcomputers," *Research Policy*, 28(5), 469–88.
- Funke, M., and R. Ruhwedel (2001a): "Export Variety and Export Performance: Empirical Evidence from East Asia," *Journal of Asian Economics*, 12(4), 493–505.
- (2001b): "Product Variety and Economic Growth: Empirical Evidence for the OECD Countries," *IMF Staff Papers*, 48(2), 225–42.
- (2002): "Export Variety and Export Performance: Empirical Evidence for the OECD Countries," *Weltwirtschaftliches Archiv/Review of World Economics*, 138(1), 97–114.
- Gabszewicz, J. J., and J. F. Thisse (1980): "Entry (and Exit) in a Differentiated Industry," *Journal of Economic Theory*, 22(2), 327–38.
- Goldin, C., and L. F. Katz (2000): "Education and Income in the Early Twentieth Century: Evidence from the Prairies," *Journal of Economic History*, 60(3), 782–818.
- (2001): "The Legacy of U.S. Educational Leadership: Notes on Distribution and Economic Growth in the 20th Century," *American Economic Review*, 91(2), 18–23.
- Goolsbee, A., and A. Petrin (2001): *The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable Television*. NBER working paper series 8317.
- Greenwood, J., and G. Uysal (2004): *New Goods and the Transition to a New Economy*. NBER Working Paper 10793.
- Grossman, G. M., and E. Helpman (1989): "Product Development and International Trade," *Journal of Political Economy*, 97(6), 1261–83.
- (1991a): *Innovation and growth in the global economy*. MIT Press, Cambridge and London.
- (1991b): "Quality Ladders and Product Cycles," *Quarterly Journal of Economics*, 106(2), 557–86.
- (1991c): "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, 58(1), 43–61.
- Hanoch, G. (1971): "CRESH Production Functions," *Econometrica*, 39(5), 695–712.
- Hausman, J. A. (1981): "Exact Consumer's Surplus and Deadweight Loss," *American Economic Review*, 71(4), 662–76.
- (1997): "Valuation of New Goods under Perfect and Imperfect Competition," in *The economics of new goods*, ed. by T. F. Bresnahan, and R. J. Gordon, pp. 209–37. University of Chicago Press, Chicago and London.
- (2003): "Sources of Bias and Solutions to Bias in the Consumer Price Index," *Journal of Economic Perspectives*, 17(1), 23–44.

- Helpman, E., and P. Krugman (1985): *Market structure and foreign trade*. MIT Press, Cambridge.
- Heston, A., R. Summers, and B. Aten (2002): *Penn World Table Version 6.1*. Center for International Comparisons at the University of Pennsylvania (CICUP).
- Hicks, J. R. (1940): "The valuation of social income," *Economica*, 7(26), 105–124.
- Hotelling, H. (1929): "Stability in competition," *Economic Journal*, 39(153), 41–57.
- Howitt, P. (1999): "Steady Endogenous Growth with Population and R&D Inputs Growing," *Journal of Political Economy*, 107(4), 715–30.
- Hummels, D., and P. J. Klenow (2002): *The Variety and Quality of a Nation's Trade*. NBER Working Paper 8712.
- Jaffe, A. B. (1986): "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value," *American Economic Review*, 76(5), 984–1001.
- Jefferson, G. H., and T. G. Rawski (1994): "Enterprise Reform in Chinese Industry," *Journal of Economic Perspectives*, 8(2), 47–70.
- Jones, C. I. (1995a): "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 103(4), 759–84.
- (1995b): "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics*, 110(2), 495–525.
- (1998): *Introduction to Economic Growth*. W. W. Norton & Company, New York.
- (1999): "Growth: With or Without Scale Effects?," *American Economic Review*, 89(2), 139–44.
- (2002): "Sources of U.S. Economic Growth in a World of Ideas," *American Economic Review*, 92(1), 220–39.
- Judd, K. L. (1985a): "The Law of Large Numbers with a Continuum of IID Random Variables," *Journal of Economic Theory*, 35(1), 19–25.
- (1985b): "On the Performance of Patents," *Econometrica*, 53(3), 567–85.
- Kaldor, N. (1935): "Market imperfection and excess capacity," *Economica*, 2(5), 33–50.
- Kamien, M. I., and N. L. Schwartz (1991): *Dynamic optimization: The calculus of variations and optimal control in economics and management*. North-Holland; distributed in the U.S. and Canada by Elsevier Science, New York, Amsterdam; London and Tokyo.
- Kapur, J. N., and H. Kesavan (1992): *Entropy Optimization Principles with Applications*. Academic Press, San Diego.
- Kerr, I. J. (1995): "Colonialism and Technological Choice: The Case of the Railways of India," *Itinerario*, 19(2), 91–111.
- Klenow, P. J. (2003): "Measuring Consumption Growth: The Impact of New and Better Products," *Federal Reserve Bank of Minneapolis Quarterly Review*, 27(1), 10–23.
- Klundert, T. v. d., and S. Smulders (2001): "Imperfect Competition, Concentration and Growth," in *Growth theory in historical perspective: Selected essays of Theo van de Klundert*, pp. 201–26. Elgar; distributed by American International Distribution Corporation, Williston, Vt., Cheltenham, U.K. and Northampton, Mass.

- Koopmans, T. C. (1965): "On the Concept of Optimal Economic Growth," in *The Econometric Approach to Development Planning*. North-Holland, Amsterdam.
- Kortum, S. S. (1997): "Research, Patenting, and Technological Change," *Econometrica*, 65(6), 1389–1419.
- Kremer, M. (1993): "Population Growth and Technological Change: One Million B.C. to 1990," *Quarterly Journal of Economics*, 108(3), 681–716.
- Krugman, P. R. (1979): "Increasing Returns, Monopolistic Competition, and International Trade," *Journal of International Economics*, 9(4), 469–79.
- (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70(5), 950–59.
- (1991): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483–99.
- Krugman, P. R., and A. J. Venables (1995): "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, 110(4), 857–80.
- Laitner, J. (2000): "Earnings within Education Groups and Overall Productivity Growth," *Journal of Political Economy*, 108(4), 807–32.
- Lancaster, K. (1966): "A new approach to consumer theory," *Journal of Political Economy*, 74(2), 132–157.
- (1975): "Socially Optimal Product Differentiation," *American Economic Review*, 65(4), 567–85.
- (1979): *Variety, equity, and efficiency*. Columbia University Press, New York.
- Landes, D. S. (1998): *The wealth and poverty of nations: Why some are so rich and some so poor*. Norton, New York and London.
- Laspeyres, E. (1871): "Die Berechnung einer mittleren Waarenpreissteigerung," *Jahrbücher für Nationalökonomie und Statistik*, 16, 296–314.
- Leibenstein, H. (1950): "Bandwagon, snob and Veblen effects in the theory of consumers," *Quarterly Journal of Economics*, 65, 183–207.
- Li, C. W. (2000): "Endogenous vs. Semi-endogenous Growth in a Two-R&D-Sector Model," *Economic Journal*, 110(462), C109–22.
- Lucas, R. E. (1988): "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22(1), 3–42.
- Malerba, F., and L. Orsenigo (1995): "Schumpeterian Patterns of Innovation," *Cambridge Journal of Economics*. February, 19(1), 47–65.
- Malerba, F., L. Orsenigo, and P. Peretto (1997): "Persistence of Innovative Activities, Sectoral Patterns of Innovation and International Technological Specialization," *International Journal of Industrial Organization*, 15(6), 801–26.
- Malliari, A. G., and W. A. Brock (1982): *Stochastic methods in economics and finance*. North-Holland, New York.
- Marryat, J. (1857): *A history of pottery and porcelain, mediaeval and modern*. John Murray, London, 2d edn.
- Maurseth, P. B., and B. Verspagen (2002): "Knowledge Spillovers in Europe: A Patent Citations Analysis," *Scandinavian Journal of Economics*, 104(4), 531–45.
- Metcalf, J. S. (1998): *Evolutionary economics and creative destruction*. Routledge, London.
- (1999): *Competition, Fisher's Principle and Increasing Returns in the Selection Process*. Elgar, Cheltenham, U.K. and Northampton, Mass.

- Mokyr, J. (1990): *The lever of riches: Technological creativity and economic progress*. Oxford University Press, Oxford; New York; Toronto and Melbourne.
- (1993): "The New Economic History and the Industrial Revolution," in *The British Industrial Revolution: An economic perspective.*, pp. 1–131. Westview Press, Boulder and Oxford.
- (2002): *The gifts of Athena: Historical origins of the knowledge economy*. Princeton University Press, Princeton and Oxford.
- Mossay, P. (2003): "Increasing Returns and Heterogeneity in a Spatial Economy," *Regional Science and Urban Economics*, 33(4), 419–44.
- Mussard, S., F. Seyte, and M. Terraza (2003): "Decomposition of Gini and the Generalized Entropy Inequality Measures," *Economics Bulletin*, 4(7), 1–6.
- Nehring, K., and C. Puppe (2002): "A Theory of Diversity," *Econometrica*, 70(3), 1155–98.
- Nelson, R. R., and S. G. Winter (1982): *An evolutionary theory of economic change*. Harvard University Press, Cambridge, Mass.
- Nevo, A. (2003): "New Products, Quality Changes, and Welfare Measures Computed from Estimated Demand Systems," *Review of Economics and Statistics*, 85(2), 266–75.
- Nordhaus, W. D. (1997): "Do Real-Output and Real-Wage Measures Capture Reality? The History of Lighting Suggests Not," in *The economics of new goods*, ed. by T. F. Bresnahan, and R. J. Gordon, pp. 29–66. University of Chicago Press, Chicago and London.
- OECD (2002a): *International Trade by Commodity Statistics (ITCS); SITC rev. 2; historical series*. Paris.
- (2002b): *International Trade by Commodity Statistics (ITCS); SITC rev.3*. Paris.
- (2004): *Main Science and Technology Indicators (MSTI)*. Paris.
- O'Rourke, K. H., and J. G. Williamson (2000): *When Did Globalization Begin?* NBER Working Paper 7632.
- (2002): "When Did Globalisation Begin?," *European Review of Economic History*, 6(1), 23–50.
- Paassche, H. (1874): "Über die Preisentwicklung der letzten Jahre nach den Hamburger Borsennotirungen," *Jahrbücher für Nationalökonomie und Statistik*, 12, 168–178.
- Paine, C. (1936): "Rationalisation and the theory of excess capacity," *Economica*, 3(9), 46–60.
- Peretto, P. F. (1998a): "Technological Change and Population Growth," *Journal of Economic Growth*, 3(4), 283–311.
- (1998b): "Technological Change, Market Rivalry, and the Evolution of the Capitalist Engine of Growth," *Journal of Economic Growth*, 3(1), 53–80.
- (2003a): "Endogenous market structure and the growth and welfare effects of economic integration," *Journal of International Economics*, 60, 177–201.
- (2003b): "Fiscal Policy and Long-Run Growth in R&D-Based Models with Endogenous Market Structure," *Journal of Economic Growth*, 8(3), 325–47.
- Peretto, P. F., and S. Smulders (2002): "Technological Distance, Growth and Scale Effects," *Economic Journal*, 112(481), 603–24.
- Perkins, D. H. (1994): "Completing China's Move to the Market," *Journal of Economic Perspectives*, 8(2), 23–46.

- Petrin, A. (2001): *Quantifying the Benefits of New Products: The Case of the Minivan*. NBER working paper series 8227.
- (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110(4), 705–29.
- Pettengill, J. S. (1979): "Monopolistic Competition and Optimum Product Diversity: Comment," *American Economic Review*, 69(5), 957–60.
- Pomeranz, K. (2000): *The Great Divergence: China, Europe, and the making of the modern world economy*. Princeton University Press, Princeton and Oxford.
- Puga, D., and A. J. Venables (1999): "Agglomeration and Economic Development: Import Substitution vs. Trade Liberalisation," *Economic Journal*, 109(455), 292–311.
- Rae, J. (1905 [1834]): *The sociological theory of capital [The new principles of political economy]*. MacMillan, London.
- Ram, R. (1990): "Educational Expansion and Schooling Inequality: International Evidence and Some Implications," *Review of Economics and Statistics*, 72(2), 266–74.
- Ramsey, F. (1928): "A Mathematical Theory of Saving," *Economic Journal*, 38(December), 543–559.
- Reinstaller, A., and B. Sanditov (2003): *Social norms and equality of opportunity in conspicuous consumption: on the diffusion of consumer good innovation*. MERIT Research Memorandum 2003-014.
- Rivera Batiz, L. A., and P. M. Romer (1991): "Economic Integration and Endogenous Growth," *Quarterly Journal of Economics*, 106(2), 531–55.
- (1994): "Economic Integration and Endogenous Growth: An Addendum," *Quarterly Journal of Economics*, 109(1), 307–08.
- Robinson, J. (1934): "What is imperfect competition?," *Quarterly Journal of Economics*, 49(1), 104–120.
- (1950 [1933]): *Economics of imperfect competition*. Macmillan and co., London.
- Romer, P. M. (1987): "Growth Based on Increasing Returns Due to Specialization," *American Economic Review*, 77(2), 56–62.
- (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98(5), S71–S102.
- (1994): "New Goods, Old Theory, and the Welfare Costs of Trade Restrictions," *Journal of Development Economics*, 43(1), 5–38.
- Rosen, S. (1976): "A Theory of Life Earnings," *Journal of Political Economy*, 84(4), 67.
- Rutherford, T. F., and D. G. Tarr (2002): "Trade Liberalization, Product Variety and Growth in a Small Open Economy: A Quantitative Assessment," *Journal of International Economics*, 56(2), 247–72.
- Salop, S. C. (1979): "Monopolistic Competition with Outside Goods," *Bell Journal of Economics*, 10(1), 141–56.
- Sato, K. (1976): "The Ideal Log-Change Index Number," *Review of Economics and Statistics*, 58(2), 223–28.
- Sattinger, M. (1993): "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, 31(2), 831–80.
- Schankerman, M. (2002): "Idiosyncratic and Common Shocks to Investment Decisions," *Economic Journal*, 112(482), 766–85.

- Scherer, F. M. (1998): "The Size Distribution of Profits from Innovation," *Annales d'Economie et de Statistique*, 0(49-50), 495-516.
- Scherer, F. M., D. Harhoff, and J. Kukies (2000): "Uncertainty and the Size Distribution of Rewards from Innovation," *Journal of Evolutionary Economics*, 10(1-2), 175-200.
- Schumpeter, J. A. (1934): *The Theory of Economic Development*. Harvard University Press, Cambridge MA.
- (1943): *Capitalism, Socialism and Democracy*. George Allen & Unwin, London.
- Schwartz, B. (2004): *The paradox of choice: Why more is less*. HarperCollins, Ecco, New York.
- Segerstrom, P. S. (1998): "Endogenous Growth without Scale Effects," *American Economic Review*, 88(5), 1290-1310.
- (2000): "The Long-Run Growth Effects of R&D Subsidies," *Journal of Economic Growth*, 5(3), 277-305.
- Shaked, A., and J. Sutton (1982): "Relaxing Price Competition through Product Differentiation," *Review of Economic Studies*, 49(1), 3-13.
- (1990): "Multiproduct Firms and Market Structure," *RAND Journal of Economics*, 21(1), 45-62.
- Shannon, C. E. (1948): "A mathematical theory of communication," *The Bell System Technical Journal*, 27, 379-423, 623-656.
- Spence, M. (1976): "Product Selection, Fixed Costs, and Monopolistic Competition," *Review of Economic Studies*, 43(2), 217-35.
- Sraffa, P. (1926): "The laws of returns under competitive conditions," *Economic Journal*, 36(144), 535-550.
- Statistics Europe (2001): *EUROPROMS, European production and market statistics*. Brussels.
- Stokey, N. L. (1988): "Learning by Doing and the Introduction of New Goods," *Journal of Political Economy*, 96(4), 701-17.
- Straathof, B. (2002): *Micro-uncertainty and growth*. MERIT Research Memorandum 2002-01.
- Sutton, J. (1986): "Vertical Product Differentiation: Some Basic Themes," *American Economic Review*, 76(2), 393-98.
- Tang, P. J. G., and K. Wälde (2001): "International Competition, Growth and Welfare," *European Economic Review*, 45(8), 1439-59.
- Taylor, M. S. (1994): "'Once-off' and Continuing Gains from Trade," *Review of Economic Studies*, 61(3), 589-601.
- Teulings, C. N. (1995): "The Wage Distribution in a Model of the Assignment of Skills to Jobs," *Journal of Political Economy*, 103(2), 280-315.
- Theil, H. (1965): "The information approach to demand analysis," *Econometrica*, 33(1), 67-87.
- (1967): *Economics and information theory*. North Holland, Amsterdam.
- Thompson, P. (2001): "The Microeconomics of an R&D-Based Model of Endogenous Growth," *Journal of Economic Growth*, 6(4), 263-83.
- Trajtenberg, M. (1989): "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners," *Journal of Political Economy*, 97(2), 444-79.
- Triffin, R. (1949 [1940]): *Monopolistic competition and general equilibrium theory*. Harvard University Press, Cambridge.



- United Nations Statistics Division (2004): *COMTRADE*. <http://unstats.un.org/unsd/>.
- Vartia, Y. (1976): "Ideal log-change index numbers," *The Scandinavian Journal of Statistics*, 3, 121–126.
- Veblen, T. (1921): *The theory of the leisure class: an economic study of institutions*. MacMillan, New York.
- Ventura, J. (1997): "Growth and Interdependence," *Quarterly Journal of Economics*, 112(1), 57–84.
- Weitzman, M. L. (1992): "On Diversity," *Quarterly Journal of Economics*, 107(2), 363–405.
- (1994): "Monopolistic Competition with Endogenous Specialization," *Review of Economic Studies*, 61(1), 45–56.
- (1998): "Recombinant Growth," *Quarterly Journal of Economics*, 113(2), 331–60.
- Williamson, J. G. (1996): "Globalization, Convergence, and History," *Journal of Economic History*, 56(2), 277–306.
- Wolfram Research (2004): *Stirling Number of the Second Kind*. <http://mathworld.wolfram.com/StirlingNumberoftheSecondKind.html>.
- World Bank (2003): *The 2003 World Development Indicators CD-ROM*.
- World Trade Organization (1994): *Final act embodying the results of the Uruguay round of multilateral trade negotiations*. <http://www.wto.org>.
- Young, A. (1998): "Growth without Scale Effects," *Journal of Political Economy*, 106(1), 41–63.
- (2000): "The Razor's Edge: Distortions and Incremental Reform in the People's Republic of China," *Quarterly Journal of Economics*, 115(4), 1091–1135.
- Zivot, E., and D. W. K. Andrews (1992): "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis," *Journal of Business and Economic Statistics*, 10(3), 251–70.
- Zon, A. v., E. Fortune, and T. Kronenberg (2003): *How to reap and sow as you go: A simple model of cyclical endogenous growth*. MERIT Research Memorandum 2003-26.

# Index

- Aghion, Phillipe, 37, 44  
 Amoroso-Robinson condition, 15  
 autarky, 115, 117, 153, 162  
 autocatalytic process, 9
- Böttcher, 1  
 backward integration, 127  
 Baltagi-Wu estimation, 163  
 Britain, 29
- CES, 15, 16, 80  
 Chamberlin, E.H., 12–14  
   and Kaldor, 13  
   and Robinson, 12  
 China, 186–189  
   reforms in, 188  
 classification of products, 29  
   curse of, 158–160  
   HS, 29, 171, 176, 181, 191  
   NAICS, 159  
   Prodcorn, 191  
   SITC, 29, 164, 171, 173, 176, 191  
 COLI, 80–81  
 constant elasticity of substitution, *see* CES  
 consumption  
   and household savings, 47–50  
   conspicuous, 9  
 convergence, 22–23, 186–189  
   in education, 88  
 Corn Laws, 2  
 cost of living index, *see* COLI  
 CPU, 100  
 creative destruction, 37, 42, 101  
 CRESH function, 83–84
- Diewert, Erwin, 72, 73, 75, 84, 86  
 Dinopoulos, Elias, 39, 61, 150  
 distribution of income, 120–124  
 diversity value functions, 68  
 Dixit, Avinash, *see* Dixit-Stiglitz  
 Dixit-Stiglitz, 14–16, 162  
 Drobisch, 71
- economic integration, 5, 111–127  
 economies of scale, *see* scale, returns to
- economy  
   independent national, 22  
   size of, 4–5, 117, 152–155  
 education, *see also* schooling  
   advances in, 5, 87–99  
 endogenous growth, 38, 104, 108  
   generalization of, 51–53  
 entropy, 63–64, 105–106, 108–109, 136, 139–140, 172–173  
   properties of, 64–65  
 EPO, 26  
 Ethier, Wilfred, 17  
 excess capacity, 12–14  
 exports, 21–22  
   data on, 162–164  
 externalities, 3
- factor price equalization, *see* FPE  
 Feenstra index, 73, 85–86, 152  
 Fisher index, 72, 76  
 Fisher's theorem, 5, 131–132  
 fishing-out, 4  
 fixed cost, 16, 113, 115  
 FPE, 114, 118  
 free entry, 15
- GDP  
   and population size, 22  
   and welfare, 186  
 general purpose technology, *see* GPT  
 Gini coefficient, 68  
 gold standard, 2  
 GPT, 4  
 Grossman, Gene, 17, 37, 44, 53  
 Grossman, Gene, 113
- Hausman, Jerry, 73, 152  
 Heckscher-Ohlin, *see also* FPE, 2  
 Helpman, Elhanan, 17, 37, 44, 53, 114  
 Helpman, Elhanan, 113  
 Hicks, John, 72, 152  
 Hirschman-Herfindahl index, 67  
 Hotelling, Harold, 12, 14  
 Howitt, Peter, 37, 39, 40, 44, 150  
 HS, *see* classification of products

- human capital, 8, 88, 91, 162, 166–168, 171, 175, 176, 181, 184
- idiosyncratic performance, 5–6, 131–140, 172–173
- imitation, 115, 153
  - of porcelain, 1
- imperfect substitutes, 3–4, 99–106, 173
- imports, 21–22
  - data on, 162–164
- India, 1, 186–189
  - reforms in, 188–189
- Industrial Revolution, 29, 186–187
- intellectual property right, *see* IPR
- intermediate products, 4, 114, 115, 153
- IPR, 112–113, 126
- Itô process, 137
- Jensen's inequality, 134, 135, 147
- Jones, Charles, 4, 17, 37
  - Jones critique, 37–39
  - on schooling, 89
- JPO, 26
- Judd, Kenneth, 6, 17, 162
- Kaldor, Nicholas, 13–14
- knife-edge condition, 39, 100
- knowledge spillovers, 6–7, 37–60, 99–106, 151
  - between countries, 115, 152, 186
  - between quality and variety, 39
  - during Industrial Revolution, 186
  - negative, 100, 106
  - within network, 58–59
- Kortum, Samuel, 28, 39
- Krugman, Paul, 16, 18, 113, 114
- Lancaster, Kelvin, 12
- Laspeyres index, 72, 76
- Li, Chol Won, 39, 44, 54
- lighting, price of, 9
- love-of-variety, 37
- Lucas, Robert, 89, 107
- Markov process, 133
- mergers, cross-border, 117
- MNE, 114
- Moluccas, the, 1
- monopolistic competition, 7
  - and excess capacity, 12–14
  - and product variety, 13
  - origins, 11–12
- multinational enterprise, *see* MNE
- NAICS, *see* classification of products
- Napoleonic Wars, 2
- neo-endogenous growth, 39
  - generalization of, 55–59
- network, 39
- Neumann John von, 64
- Newton, Isaac, 4
- numeraire, 40
- OECD, 2, 21
- Paassche index, 72, 76
- patents, 26–28, 112, 117
- Peretto, Pietro, 39, 44, 45, 58–59, 129
- Pettengill, John, 16
- population
  - growth, 37–60, 87, 93, 94, 112, 113, 185
  - size, 22, 23, 56, 60, 152–155, 162, 166, 175, 183, 184, 188
- population growth, 5
- porcelain, 1
- price index, 71–82, 186
- Prodcorn, *see* classification of products
- product classification, *see* classification of products
- product differentiation
  - horizontal, *see* product variety
  - vertical, 18
- product quality, 6, 8, 37–60, 69, 99–106, 132, 186
  - and dividends, 47
  - and Ramsey rule, 50
  - index, 80
- product test, 74
- product variety
  - and oligopoly, 18
  - and the standard of living, 3–4
  - and consumer welfare, 12–14
  - and monopolistic competition, 13
  - and the standard of living, 63
  - data on, 29
  - decline in, 6, 136
  - effective, 6, 63, 101, 105–106, 136, 139–140, 151, 172–173
  - index of, *see* variety index
- production
  - data on, 152
  - modelling of, 40–42
- productivity, 131
  - measurement of, 79–80
- quality, *see* product quality
- quality ladder, 37, 132
- quantity index, 71–82, 186
- R&D, *see* research
- Ramsey rule, 50, 103, 109, 138
- research, 4–5
  - data on, 24–28
  - modelling of, 42–47
- Robinson, Joan, 12, 15
- Romer, Paul, 7, 17, 53, 113, 114
- saddle path stability, 117, 127–128

- Sato-Vartia index, 72, 85  
 savings, 47–50  
 scale  
   of economy, 7, 39, 119, 151, 162, 166, 171, 175, 184, 187  
   of production, 119, 124, 153, 154, 175, 184  
   returns to, 12, 13, 16, 35, 43, 79, 111, 113, 117, 119, 153, 154  
   savings, 119  
   ‘scale effect’, 59  
 schooling, *see also* education  
   data on, 23–24  
   distribution of, 24  
   inequality, 87–99, 162  
 Schumpeter, Joseph  
   Mark I & II, 9  
 Segerstrom, Paul, 39, 40, 44  
 semi-endogenous growth, 38, 39  
   generalization of, 54–55  
 Shannon’s entropy, *see* entropy  
 Shannon, Claude, 63  
 shoulders, standing on, 4  
 SITC, *see* classification of products  
 skills, 87, 114  
   distribution of, 90  
 Smulders, Sjak, 39, 45, 58–59  
 spectacles, 9  
 spices, 1  
 spillovers, *see* knowledge spillovers  
 Straffa, Piero, 11  
 standard of living, 3–6, 22, 63, 131, 184, 187  
 Stiglitz, Joseph, *see* Dixit-Stiglitz  
 Stirling number, second kind, 158  
 Stokey, Nancy, 18  
  
 Theil, Henry, 64  
 Thompson, Peter, 39, 44, 61, 150  
 triadic patent family, 26, 35  
 TRIPS, 112  
  
 unit root test, with break, 164  
 USPTO, 26  
  
 variety, *see* product variety  
 variety index, 67–68, 74, 105–106, 108–109, 136, 139–140, 172–173  
  
 World Trade Organization, *see* WTO  
 Wright, Wilbur and Orville, 4  
 WTO, 112, 188  
  
 Young, Alwyn, 39, 44, 189



# Samenvatting

## PRODUCTVARIËTEIT EN ECONOMISCHE GROEI

### De tegengestelde effecten van schaal en eigenaardigheid

Onder economische groei wordt gewoonlijk groei in de hoeveelheid geproduceerde goederen verstaan. Dit proefschrift behandelt een ander aspect van economische groei, namelijk groei in de variëteit van consumentenproducten. Voor deze benadering is gekozen omdat het beter is om van verschillende producten voldoende te hebben, dan om van slechts één product heel veel.

Dit proefschrift levert een theoretische fundering en empirisch bewijs voor twee proposities. De eerste propositie stelt dat bevolkingsgroei, economische integratie en een hoger opleidingsniveau hebben bijgedragen aan een grotere productvariëteit in de landen van de OESO (Organisatie voor Economische Samenwerking en Ontwikkeling). Deze propositie is gebaseerd op het idee dat het uitvinden van een product alleen winstgevend is als de potentiële markt voor dat product groot genoeg is. De grootte van de potentiële markt hangt af van het aantal potentiële klanten en hun koopkracht. Het aantal klanten kan toenemen door bevolkingsgroei en door het verkrijgen van toegang tot andere economieën (economische integratie). De koopkracht hangt in belangrijke mate af van het scholingsniveau. Zonder bevolkingsgroei, economische integratie of meer scholing is er geen groei van (potentiële) markten mogelijk en verdwijnt de beloning voor het uitvinden van een nieuw product.

De tweede propositie stelt dat het bestaan van verschillen in de prestaties van bedrijfstakken een negatief effect heeft op de effectieve variëteit aan producten in OESO-landen. Deze propositie is gebaseerd op een mechanisme van variatie en selectie. Verschillende producten evolueren op verschillende manieren. Eén oorzaak hiervan is kostenbesparing. Van sommige producten zijn de kosten sterk afgenomen (denk aan vliegreizen), van andere niet (briefpost). Een andere oorzaak is kwaliteitsverbetering. Van sommige producten komt regelmatig een verbeterde variant op de markt (computers), maar van andere niet (suiker).

Dit soort type-specifieke ontwikkelingen zorgen er voor dat consumenten een groot deel van hun inkomen besteden aan een beperkt aantal producten. Men geeft veel uit aan producten waarvan de prijs en kwaliteit gunstig zijn, waardoor er minder besteed kan worden aan minder innovatieve producten. Het bestaan van verschillen in de prestaties van bedrijfstakken leidt er dus toe dat consumenten verschillende gewichten toekennen aan verschillende producten. Dit heeft een

negatieve invloed op de effectieve variëteit aan producten omdat een aantal producten niet als belangrijk worden ervaren.

Deze twee proposities suggereren een paradox: volgens de eerste propositie neemt productvariëteit toe, volgens de tweede propositie neemt productvariëteit af. In het empirisch deel van het proefschrift wordt aangetoond dat beide effecten duidelijk aanwezig zijn in de OESO, maar dat het positieve effect van de eerste propositie overheerst.

Naast theoretische en empirische resultaten op het gebied van economische groei, bevat dit proefschrift ook nieuwe formules voor het meten van inflatie en productiviteitsgroei.

# Curriculum Vitae

Sebastiaan (Bas) Michael Straathof was born in The Hague on February 16th 1976. He attended secondary education at C.S.G. Overvoorde from 1988 to 1994. In the latter year, he started a study Economics at the Erasmus University Rotterdam, where he obtained an MSc in Development Planning in 1999. After having worked for several months at the Agricultural Economics Research Institute (LEI) in The Hague, he started his PhD at the Maastricht Economic Research Institute on Innovation and Technology (MERIT) in October 2000. Between August 2004 and January 2005, he worked as a lecturer at the Department of Economics in Maastricht. Bas is currently employed as post-doc at the Department of Technology Management, Eindhoven University of Technology.